

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:00:53 ; Search time 20.9296 Seconds
(without alignments)
757.625 Million cell updates/sec

Title: US-09-627-165E-14

Sequence: 1 ARFNPIXWRLRQINSGE...VDAANDVTCTXSEPTVRIV 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summarles

Database :

1: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1980.DAT *
A.genseqp_101002.*
2: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1981.DAT *
3: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1982.DAT *
4: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1983.DAT *
5: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1984.DAT *
6: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1985.DAT *
7: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1986.DAT *
8: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1987.DAT *
9: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1988.DAT *
10: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1989.DAT *
11: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1990.DAT *
12: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1991.DAT *
13: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1992.DAT *
14: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1993.DAT *
15: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1994.DAT *
16: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1995.DAT *
17: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1996.DAT *
18: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1997.DAT *
19: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1998.DAT *
20: /SIDS1/gcgdata/genseq/genseqp -emb1/AA1999.DAT *
21: /SIDS1/gcgdata/genseq/genseqp -emb1/AA2000.DAT *
22: /SIDS1/gcgdata/genseq/genseqp -emb1/AA2001.DAT *
23: /SIDS1/gcgdata/genseq/genseqp -emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	590	99.0	120	22	AAB47096	Korean mistletoe I
2	413.5	69.5	174	22	ABB7450	Galatosea-recognis
3	408.5	68.5	174	22	AAB47097	Korean mistletoe I
4	387	64.9	564	18	AAH10021	Prepro mistletoe I
5	387	64.9	564	18	AAW90127	Mistletoe lectin P
6	386	64.8	531	20	AAV25979	Mistletoe lectin I
7	386	64.8	532	20	AAV25982	Mistletoe lectin I
8	373	62.6	533	20	AAV25973	Mistletoe lectin P
9	373	62.6	533	20	AAV25973	Mistletoe lectin P
10	373	62.6	533	20	AAV25976	Mistletoe lectin P

11	327	54.9	256	20	AAV25981
12	327	54.9	256	20	AAV25984
13	317	53.2	254	20	AAV25980
14	311	52.2	254	20	AAV25985
15	311	52.2	252	19	AAV646655
16	311	52.2	253	18	AAV606666
17	311	52.2	252	18	AAV100222
18	311	52.2	253	20	AAV959121
19	304	51.0	255	20	AAV259715
20	304	51.0	255	20	AAV25974
21	304	51.0	255	20	AAV25977
22	303	50.8	254	22	AAV47090
23	298	50.0	254	22	AAV47091
24	298	50.0	256	22	AAV47092
25	178	29.9	332	8	AAV70097
26	178	29.9	332	8	AAV70838
27	178	29.9	332	10	AAV956339
28	178	29.9	332	11	AAV085534
29	178	29.9	362	10	AAV90079
30	178	29.9	565	6	AAV50166
31	178	29.9	565	7	AAV60240
32	178	29.9	565	22	AAV78300
33	178	29.9	574	8	AAV70325
34	178	29.9	574	10	AAV94729
35	178	29.9	576	8	AAV70326
36	178	29.9	576	18	AAV25787
37	178	29.9	576	20	AAV55892
38	178	29.9	576	21	AAV785927
39	178	29.9	576	22	AAV78301
40	178	29.9	576	22	AAV78300
41	177	29.7	280	10	AAV956448
42	177	29.7	540	18	AAV25114
43	177	29.7	540	18	AAV21706
44	174	29.2	534	8	AAV70324
45	165	27.7	565	22	AAV78304

ALIGNMENTS

RESULT	1
AAAB47096	
ID	AAAB47096 standard; Protein: 120 AA.
XX	
AC	AAAB47096;
XX	
DT	16-MAY-2001 (first entry)
XX	
DE	Korean mistletoe lectin #1.
XX	
KW	isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe
KW	KML; tumour; KM-110; KML-C; KMBP; KML-110; KML-11L;
KW	heparin binding protein.
XX	
OS	Viscum album coloratum.
XX	
EH	Key
FT	Misc-difference 7 Location/Qualifiers
FT	/note= "Encoded by NTG"
FT	Misc-difference 19
FT	/note= "Encoded by TCN"
FT	Misc-difference 64
FT	/note= "Encoded by NTG"
FT	Misc-difference 111
FT	/note= "Encoded by NTG"
XX	
PN	EP1074560-A2.
XX	
PD	07-FEB-2001.
XX	
PF	27-JUL-2000; 2000EP-0402163.
XX	
PR	27-JUL-1999; 99KR-0030638.

Mistletoe lectin A
 Mistletoe lectin A
 Mistletoe lectin A
 Mistletoe lectin A
 Mistletoe rRNA protein
 Mistletoe rRNA variant
 Prepro mistletoe I
 Mistletoe ML A-chain
 Mistletoe lectin A
 Mistletoe lectin A
 Mistletoe lectin A
 A-chain isoform for
 A-chain isoform for
 A-chain isoform for
 Ricin A. Escherich
 Sequence of Ricinu
 Ricin A encoded by
 Ricin A gene product
 Ricin D. Ricinus
 Sequence of prepro
 PreproRiclin. Ricl
 Castor bean prepro
 Sequence of Ricinu
 DNA sequence of rR
 Sequence of Ricinu
 Castorbean riclin.
 Castor bean riclin
 Ricinus communis r
 Castor bean prepro
 Castor bean prepro
 Ricin agglutinin A
 Castor oil plant a
 R. communis agglut
 Sequence of Ricin
 Modified castor be

XX (MIST-) MISTLE BIOTECH CO LTD.
 PA
 XX
 PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX
 DR WPI: 2001-171044/18.
 DR N-PSDB; AAC47096.
 XX
 PT Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 PS
 XX Claim 31; Page 33; 62pp; English.
 CC
 CC The sequences given in AAB47096-97 are lectins isolated from Korean
 CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
 CC immunity and for treating tumours. The KML's are isolated from a
 CC protein fraction derived from the leaves, stems and fruits of Korean
 CC mistletoe, which is designated KM-110. One of the isolates, KML-C was
 CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
 CC lymphoma.
 CC
 XX Sequence 120 AA;

Query Match 99.0%; Score 590; DB 22; Length 120;
 Best Local Similarity 100.0%; Pred. No. 8,7e-66;
 Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARFNPIMRLRQINSSESPNNMYMLETSMGRSTQVOQSKDGIENQIRLQISAGN 60
 DB 1 ARFNPIMRLRQINSSESPNNMYMLETSMGRSTQVOQSKDGIENQIRLQISAGN 60
 QY 61 FVTXSNVRDVISSLAIMLFECSGRPFSSLDHPSPLLRSVVDANDVCTXSEPTVYR 119
 DB 61 FVTXSNVRDVISSLAIMLFECSGRPFSSLDHPSPLLRSVVDANDVCTXSEPTVYR 119

RESULT 2
 AAB79450
 ID AAB79450 standard; Protein; 551 AA.
 XX

AC AAB79450;
 DT 08-JUL-2002 (first entry)
 XX
 DE Galactose-recognising mistletoe lectin.
 XX
 KM Mistletoe; galactose-recognising mistletoe lectin; MLIIT.
 OS
 XX Viscum album.
 XX

XX Key Location/Qualifiers
 FH Misc-difference 223 /note= "Encoded by ATG"
 FT Misc-difference 251 /note= "Encoded by TTT"
 FT Misc-difference 344 /note= "Encoded by TCG"
 FT Misc-difference 380 /note= "Encoded by GCC"
 FT Misc-difference 448 /note= "Encoded by GTG"
 FT

DE10044027-A1.
 PN 14-MAR-2002.
 PD 06-SEP-2000; 2000DE-1044027.
 PF 06-SEP-2000; 2000DE-1044027.
 PR (VISC-) VISCUM AG.
 PA
 XX

PI Kleff S;
 XX
 DR WPI: 2002-316737/36.
 DR N-PSDB; ABL56947.
 XX
 PT New nucleic acid encoding preprotein of mistletoe lectin, useful as
 PT diagnostic and therapeutic agents, also encodes polypeptide -
 PS
 XX Claim 1; Fig 1; 6pp; German.
 CC
 CC The invention relates to a nucleic acid molecule (ABL56947) that encodes
 CC a preprotein (AAB79450) which, after maturation, has the biological
 CC activity of the galactose-recognising mistletoe lectin (MLIIT). The MLIIT
 CC encoding nucleic acid molecule, primers specific to it or complements of
 CC it, and encoded (oligomeric) polypeptides are useful as diagnostic and
 CC therapeutic agents.
 CC
 XX Sequence 551 AA;

Query Match 69.4%; Score 413.5; DB 23; Length 551;
 Best Local Similarity 69.7%; Pred. No. 8,3e-43;
 Matches 85; Conservative 12; Mismatches 22; Indels 3; Gaps 1;
 QY 1 ARFNPIMRLRQINSSESPNNMYMLETSMGRSTQVOQSKDGIENQIRLQISAGN 60
 DB 200 ARFNPIMRLRQINSSESPNNMYMLETSMGRSTQVOQSKDGIENQIRLQISAGN 259
 QY 61 FVTXSNVRDVISSLAIMLFECSGRPFSSLDHPSPLLRSVVD--ANDVCTXSEPTVYR 117
 DB 260 FVTXSNVRDVISSLAIMLFECSGRPFSSLDHPSPLLRSVVD--ANDVCTXSEPTVYR 319
 QY 118 IV 119
 DB 320 IV 321

RESULT 3
 AAB47097
 ID AAB47097 standard; Protein; 174 AA.
 XX

AC AAB47097;
 DT 16-MAR-2001 (first entry)
 XX
 DE Korean mistletoe lectin #2.
 XX
 KM Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
 KM KML; tumour; KM-110; KML-C; KMLBP; KML-T10; KML-T1L;
 KM heparin binding protein.
 OS
 XX Viscum album coloratum.
 XX

XX EP1074560-A2.
 PN 07-FEB-2001.
 PD 27-JUL-2000; 2000EP-0402168.
 PF 27-JUL-1999; 99KR-0030638.
 PR
 XX (MIST-) MISTLE BIOTECH CO LTD.
 PA
 XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 PI WPI: 2001-171044/18.
 DR N-PSDB; AAC85479.
 DR

Novel lectin proteins isolated from Korean mistletoe, useful for
 enhancing immunity and effectuating anti-tumoral activity -
 PS
 XX Claim 32; Page 34; 62pp; English.
 XX

AY25979	standard; Protein; 531 AA.
AA25979	
AA25979	
18-Oct-1999	(first entry)
Mistletoe lectin I protein fragment.	
Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity; ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response; lymphokine-producing macrophage; uncontrolled cell growth; treatment; cancer; cytotoxicity; antigen; isoform; lectin I.	
Viscum album.	
DE19804210-A1.	
12-AUG-1999.	
03-FEB-1998;	98DE-1004210.
03-FEB-1998;	98DE-1004210.
(BIOS-) BIOSYN ARZNEIMITTEL GMBH.	
Morris P, Stiefel T, Voelter W, Welters P;	
WPI; 1999-445335/38.	
N-PSDB; AA209103.	
Preparation of mistletoe lectins in heterologous systems, particularly for use as anticancer agents and immunostimulants	
Claim 7; Fig 1B; 78pp; German.	
This invention describes a novel mistletoe lectin (I) and its fragments which have antitumour and immunostimulatory activity. The A-chain (MLA) of the mistletoe lectin binds to, and inactivates, the 28S subunit of ribosomes. Non-cytotoxic forms of (I) activate T-cell and lymphokine-producing macrophages, so stimulate immunity. (I) and its fragments are used to treat uncontrolled cell growth (particularly cancers) and if they lack cytotoxicity, to increase the strength of the immune response, particularly to a co-administered antigen (tumour-associated, bacterial or viral). The method allows production of mistletoe lectin, and its individual chains, in many different isoforms and on a large scale, at any time of the year. Recombinant products are free from toxins present in natural mistletoe extracts. This sequence represents a mistletoe lectin I protein fragment.	
Sequence	531 AA;
Query Match	64.8%; Score 386; DB 20; Length 531;
Best Local Similarity	67.2%; Pred. No. 2.2e-39;
Matches	80; Conservative 12; Mismatches 25; Indels 2; Gaps 1
1	ARFNPILRLARQINSGSSPPNMTMLEETSMGRSTOVQOSKDIPTMTIRLOISAGN 60
167	ARFNPILRLARQINSGSFLPDVYMLEETSMGQOSTOVQSTBDVFNNPILRALPPGN 226
61	FVTXSNVADVISLAIMLEFCSGRPFSSLDHPSPLLRSVDAANDVTCTXSEPTVRIY 119
227	FVTILNVDVIAISLAIMLEFVCGERRSSSDVRYWPLVIRVI--ADVTCSASBPYRIY 283
RESULT 7	
AY25982	standard; Protein; 532 AA.
AY25982	
18-Oct-1999	(first entry)

[illegible]


```

OS Viscum album.
XX Key Location/Qualifiers
FH Misc-difference 15
FT /label= Asp, Glu
FT Misc-difference 63
FT /label= Gly, Gln
FT Misc-difference 66
FT /label= Ile, Val
FT Misc-difference 75
FT /label= Leu, Ala
FT Misc-difference 107
FT /label= Asp, Arg or none
FT Misc-difference 113
FT /label= Asn, Thr
FT Misc-difference 117
FT /label= Pro, Thr
FT Misc-difference 134
FT /label= Asp, Glu
FT Misc-difference 141
FT /label= Ser, Thr
FT Misc-difference 145
FT /label= Phe, Tyr
FT Misc-difference 152
FT /label= Thr, Ala
FT Misc-difference 177
FT /label= Ala, Tyr
FT Misc-difference 180
FT /label= Tyr, Asp
FT Misc-difference 185
FT /label= Ala, Glu
FT Misc-difference 191
FT /label= Val, Met
FT Misc-difference 219
FT /label= Ile, Phe
FT Misc-difference 224
FT /label= Pro, Ser
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
FT Misc-difference 287
FT /label= Asn, Ser
FT Misc-difference 290
FT /label= Cys, Arg
FT Misc-difference 325
FT /label= Gly, Asn
FT Misc-difference 364
FT /label= Gly, Asp
FT Misc-difference 426
FT /label= Gly, Gln
FT Misc-difference 435
FT /label= Val, Asp
FT Misc-difference 439
FT /label= Gln, Lys
FT Misc-difference 442
FT /label= Gly or none
FT Misc-difference 443
FT /label= Arg, Lys
FT Misc-difference 464
FT /label= Cys, Ser, Val
FT Misc-difference 480
FT /label= Ala, Gly
FT Misc-difference 481
FT /label= Gly, Ala
FT Misc-difference 483
FT /label= Ser, Gly
FT Misc-difference 484
FT /label= Gly, Ser
FT Misc-difference 493
FT /label= Gly, Tyr

```

```

FT Misc-difference 500
FT /label= Asn, Ser, Thr, Lys
FT Misc-difference 501
FT /label= Ser, Gly
FT Misc-difference 502
FT /label= Leu, Pro
FT Misc-difference 503
FT /label= Ala, Met
FT Misc-difference 504
FT /label= Met, Val
FT Misc-difference 533
FT /label= Pro, Phe
XX
XX DE19804210-A1.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Walters P;
XX MPI; 1999-445335/38.
XX
XX Claim 1; Page 25-26; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
XX which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX CC fragments are used to treat uncontrolled cell growth (particularly
XX CC cancers) and if they lack cytotoxicity, to increase the strength of the
XX CC immune response, particularly to a co-administered antigen
XX CC (tumour-associated, bacterial or viral). The method allows production of
XX CC mistletoe lectin, and its individual chains. In many different isoforms
XX CC and on a large scale, at any time of the year. Recombinant products are
XX CC free from toxins present in natural mistletoe extracts. This sequence
XX CC represents a consensus sequence of the mistletoe lectin described in the
XX CC specification.
XX
XX SQ Sequence 533 AA:

```

```

Query Match 62.6%; Score 373; DB 20; Length 533;
Best Local Similarity 65.5%; Pred. No. 9.3e-38;
Matches 78; Conservative 10; Mismatches 29; Indels 2; Gaps 1;

OY 1 ARFNPIXRRLRQINSGEPPNMYMLEETSMGROSTVOOQSKDFMTQIRLQISGN 60
DB 168 ARFNPIXRRLRQINSGEPPNMYMLEETSMGROSTVOOQSKDFMTQIRLQISGN 227
OY 61 FVTXSNRVDVVISLAIMFECSGRPPSSLDHPSPLLSVDAANDVCTXSEPTVRIV 119
DB 228 FVTXSNRVDVVISLAIMFECSGRPPSSLDHPSPLLSVDAANDVCTXSEPTVRIV 284

```

```

RESULT 9
AAY25973
ID AAY25973 standard; protein; 533 AA.
XX
XX AAY25973;
AC 18-OCT-1999 (first entry)
XX
XX DE Mistletoe lectin protein consensus sequence 2.
XX
XX KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;

```

KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform.
 XX
 OS Viscum album.
 XX Key Location/Qualifiers
 FH Misc-difference 15
 FT /label= Asp, Glu
 FT Misc-difference 63
 FT /label= Gly, Gln
 FT Misc-difference 66
 FT /label= Ile, Val
 FT Misc-difference 75
 FT /label= Leu, Ala
 FT Misc-difference 107
 FT /label= Asp, Arg, none
 FT Misc-difference 113
 FT /label= Asn, Thr
 FT Misc-difference 117
 FT /label= Pro, Thr
 FT Misc-difference 134
 FT /label= Asp, Glu
 FT Misc-difference 141
 FT /label= Ser, Thr
 FT Misc-difference 145
 FT /label= Phe, Tyr
 FT Misc-difference 152
 FT /label= Thr, Ala
 FT Misc-difference 177
 FT /label= Ala, Tyr
 FT Misc-difference 180
 FT /label= Tyr, Asp
 FT Misc-difference 185
 FT /label= Ala, Glu
 FT Misc-difference 191
 FT /label= Val, Met
 FT Misc-difference 219
 FT /label= Ile, Phe
 FT Misc-difference 224
 FT /label= Phe, Ser
 FT Misc-difference 225
 FT /label= Pro, Thr
 FT Misc-difference 232
 FT /label= Thr, Ser
 FT Misc-difference 236
 FT /label= Asp, Ser
 FT Misc-difference 287
 FT /label= Asn, Ser
 FT Misc-difference 290
 FT /label= Cys, Arg
 FT Misc-difference 325
 FT /label= Gly, Asn
 FT Misc-difference 364
 FT /label= Gly, Asp
 FT Misc-difference 427
 FT /label= Gly, Gln
 FT Misc-difference 435
 FT /label= Val, Asp
 FT Misc-difference 439
 FT /label= Gln, Lys
 FT Misc-difference 442
 FT /label= Gly, none
 FT Misc-difference 443
 FT /label= Arg, Lys
 FT Misc-difference 464
 FT /label= Cys, Ser, Val
 FT Misc-difference 480
 FT /label= Ala, Gly
 FT Misc-difference 481
 FT /label= Gly, Ala
 FT Misc-difference 483
 FT /label= Ser, Gly

FT Misc-difference 484
 FT /label= Gly, Ser
 FT Misc-difference 493
 FT /label= Gly, Tyr
 FT Misc-difference 500
 FT /label= Asn, Ser, Thr, Lys
 FT Misc-difference 501
 FT /label= Ser, Gly
 FT Misc-difference 502
 FT /label= Leu, Pro
 FT Misc-difference 503
 FT /label= Ala, Met
 FT Misc-difference 504
 FT /label= Met, Val
 FT Misc-difference 533
 FT /label= Pro, Phe
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 DR WPI; 1999-445335/38.
 XX
 PS Preparation of mistletoe lectins in heterologous systems,
 particularly for use as anticancer agents and immunostimulants
 Claim 4; Page 28-29; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of the mistletoe lectin described in the
 CC specification.
 CC
 SQ Sequence 533 AA;
 XX
 Query Match 62.6%; Score 373; DB 20; Length 533;
 Best Local Similarity 65.5%; Pred. No. 9, 3e-38;
 Matches 78; Conservative 10; Mismatches 29; Indels 2; Gaps 1;
 QY 1 ARENPXWRLRQINSSESPPMWLELTSNGROSTGVQOSKIGFTQIRLQISAGN 60
 ||||| :
 DB 168 ARFNPILMRXRXKINSXGSPFLPDXYMELFETSWGOSTGVQSHDGVFNPNPRLAIXGN 227
 ||||| :
 QY 61 FVYXSNRDVYSSLAIMLFECGSRPSSLDHPSPLLRSVYDANVCTXSEPTVRIY 119
 ||| ||| :
 DB 228 FVTLXNVRYIASIALIMLFVCGERPSSDVRYWPLVIRPVI--ADDVTSASEPTVRIY 284
 ||| ||| :
 RESULT 10
 AAY25976
 ID AAY25976 standard; protein; 533 AA.
 XX
 AC AAY25976;
 XX
 DT 18-OCT-1999 (first entry)

XX AAY25981;
 AC 18-OCT-1999 (first entry)
 XX
 DT
 XX
 DE Mistletoe lectin A2 protein fragment.
 XX
 KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KM cancer; cytotoxicity; antigen; isoform; lectin A2.
 XX
 OS Viscum album.
 XX
 PN DEL9804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PE 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 DR WPI; 1999-445335/38.
 DR N-PSDB; AAZ09105.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Claim 8; Fig 3B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of the mistletoe lectin A2 protein.
 XX
 SQ Sequence 256 AA;
 Query Match 54.9%; Score 327; DB 20; Length 256;
 Best Local Similarity 73.9%; Pred. No. 2e-32;
 Matches 65; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARENPIMRLRQINSSESPPMYMLETSMGRSTOVQOSKDGIFNTQIRLOISAGN 60
 DB 169 ARENPIMRLRQINSSESPPMYMLETSMGRSTOVQOSKDGIFNTQIRLOISAGN 228
 QY 61 FVTXSNVRDYISLAIMLFECGSRPSS 88
 DB 229 FVTLSNVRDYISLAIMLFVCGSRPSS 256
 RESULT 12
 AAY25984
 ID AAY25984 standard; Protein; 256 AA.
 XX
 AC AAY25984;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Mistletoe lectin A2 (variant) protein fragment.

KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KM cancer; cytotoxicity; antigen; isoform; lectin A2.
 XX
 OS Viscum album.
 XX
 PN DEL9804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PE 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 DR WPI; 1999-445335/38.
 DR N-PSDB; AAZ09108.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure; Fig 6B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A2 protein variant.
 XX
 SQ Sequence 256 AA;
 Query Match 54.9%; Score 327; DB 20; Length 256;
 Best Local Similarity 73.9%; Pred. No. 2e-32;
 Matches 65; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARENPIMRLRQINSSESPPMYMLETSMGRSTOVQOSKDGIFNTQIRLOISAGN 60
 DB 169 ARENPIMRLRQINSSESPPMYMLETSMGRSTOVQOSKDGIFNTQIRLOISAGN 228
 QY 61 FVTXSNVRDYISLAIMLFECGSRPSS 88
 DB 229 FVTLSNVRDYISLAIMLFVCGSRPSS 256
 RESULT 13
 AAY25980
 ID AAY25980 standard; Protein; 254 AA.
 XX
 AC AAY25980;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Mistletoe lectin A1 protein fragment.
 XX
 KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KM cancer; cytotoxicity; antigen; isoform; lectin A1.
 XX
 OS Viscum album.

PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOSYN ARZNEIMITTEL GMBH.
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX
 DR WPI: 1999-445335/38.
 XX
 DR N-PSDB; AAZ09104.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Disclosure: Fig 2B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of the mistletoe lectin A1 protein.
 XX
 SQ Sequence 254 AA:
 SQ
 Query Match 53.2%; Score 317; DB 20; Length 254;
 Best Local Similarity 71.6%; Pred. No. 3.5e-31;
 Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 Db 167 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 QY 61 FVYTSNVRDVISSIAIMLFECGSRPSS 88
 Db 227 FVTLTNVRDVISSIAIMLFECGSRPSS 254
 RESULT 14
 AAY25983
 ID AAY25983 standard; Protein: 254 AA.
 XX
 AC AAY25983;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Mistletoe lectin A1 (variant) protein fragment.
 XX
 OS Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribosome; 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A1.
 XX
 OS Viscum album.
 OS
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX

XX
 PA (BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX
 DR WPI: 1999-445335/38.
 XX
 DR N-PSDB; AAZ09107.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Disclosure: Fig 5B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A1 protein variant.
 XX
 SQ Sequence 254 AA:
 SQ
 Query Match 53.2%; Score 317; DB 20; Length 254;
 Best Local Similarity 71.6%; Pred. No. 3.5e-31;
 Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 Db 167 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 QY 61 FVYTSNVRDVISSIAIMLFECGSRPSS 88
 Db 227 FVTLTNVRDVISSIAIMLFECGSRPSS 254
 RESULT 15
 AAW64659
 ID AAW64659 standard; Protein: 252 AA.
 XX
 AC AAW64659;
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Mistletoe rMLA protein.
 XX
 OS Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation.
 XX
 OS Viscum album.
 OS
 PN WO9829540-A2.
 XX
 PD 09-JUL-1998.
 XX
 PF 02-JAN-1998; 98WO-EP00009.
 XX
 PR 02-JAN-1997; 97EP-0100012.
 XX
 PA (BRAI-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
 XX
 PI Eck J, Schmidt A, Zinke H;
 XX
 DR WPI: 1998-388122/33.
 XX
 DR N-PSDB; AAV51341.
 XX

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:25:21 : Search time 7.77387 Seconds

(without alignments)
450.397 Million cell updates/sec

Title: US-09-627-165E-14

Perfect score: 596

Sequence: 1 ARFNPXWRLRRQINSSESS.....VVDANDVCTXSEPTVRIV 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents-AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCUTS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	387	64.9	564	4	US-08-776-059-35 Sequence 35, Appl
2	311	52.2	235	4	US-08-776-059-39 Sequence 39, Appl
3	311	52.2	235	4	US-08-776-059-31 Sequence 31, Appl
4	177	29.7	540	1	US-08-378-761A-77 Sequence 77, Appl
5	177	29.7	540	1	US-08-485-286-77 Sequence 77, Appl
6	159	26.7	534	2	US-08-356-786-10 Sequence 10, Appl
7	138	23.2	267	1	US-07-901-707-1 Sequence 1, Appl
8	138	23.2	267	1	US-07-988-430-1 Sequence 1, Appl
9	138	23.2	267	1	US-08-218-303-16 Sequence 16, Appl
10	138	23.2	267	1	US-08-425-336-1 Sequence 1, Appl
11	138	23.2	267	1	US-08-488-113B-1 Sequence 1, Appl
12	138	23.2	267	1	US-08-477-484B-1 Sequence 1, Appl
13	138	23.2	267	2	US-08-646-360-1 Sequence 61, Appl
14	138	23.2	267	2	US-08-839-793D-61 Sequence 61, Appl
15	138	23.2	267	4	US-09-136-389-1 Sequence 1, Appl
16	138	23.2	267	4	US-09-610-838-1 Sequence 1, Appl
17	138	23.2	267	5	PCT-US92-09487-1 Sequence 1, Appl
18	138	23.2	267	5	PCT-US92-09487-1 Sequence 1, Appl
19	138	23.2	268	2	US-08-356-786-8 Sequence 8, Appl
20	136	22.8	290	1	US-08-378-761A-27 Sequence 27, Appl
21	136	22.8	290	1	US-08-485-286-27 Sequence 27, Appl
22	136	22.8	290	6	5248606-4 Patent No. 5248606
23	98	16.4	250	1	US-08-378-761A-71 Sequence 71, Appl
24	98	16.4	250	1	US-08-485-286-71 Sequence 71, Appl
25	92	15.4	263	1	US-07-901-707-7 Sequence 7, Appl
26	92	15.4	263	1	US-07-988-430-7 Sequence 7, Appl
27	92	15.4	263	1	US-08-425-336-7 Sequence 7, Appl

28	92	15.4	263	1	US-08-488-113B-7 Sequence 7, Appl
29	92	15.4	263	1	US-08-477-484B-7 Sequence 7, Appl
30	92	15.4	263	2	US-08-646-360-7 Sequence 7, Appl
31	92	15.4	263	4	US-08-839-765-7 Sequence 7, Appl
32	92	15.4	263	4	US-09-136-389-7 Sequence 7, Appl
33	92	15.4	263	4	US-09-610-838-7 Sequence 7, Appl
34	92	15.4	263	5	PCT-US92-09487-7 Sequence 7, Appl
35	83.5	14.0	282	1	US-08-324-301-15 Sequence 15, Appl
36	82.5	13.8	250	1	US-08-378-761A-78 Sequence 78, Appl
37	82.5	13.8	250	1	US-08-485-286-78 Sequence 78, Appl
38	79	13.3	261	1	US-07-901-707-9 Sequence 9, Appl
39	79	13.3	261	1	US-07-988-430-9 Sequence 9, Appl
40	79	13.3	261	1	US-08-425-336-9 Sequence 9, Appl
41	79	13.3	261	1	US-08-378-761A-79 Sequence 79, Appl
42	79	13.3	261	1	US-08-485-286-79 Sequence 79, Appl
43	79	13.3	261	1	US-08-488-113B-9 Sequence 9, Appl
44	79	13.3	261	1	US-08-477-484B-9 Sequence 9, Appl
45	79	13.3	261	2	US-08-646-360-9 Sequence 9, Appl

ALIGNMENTS

```
RESULT 1
US-08-776-059-35
: Sequence 35, Application US/08776059B
: Patent No. 6271368
: GENERAL INFORMATION:
: APPLICANT: LENTZEN, Hans
: APPLICANT: ECK, Jurgen
: APPLICANT: BAUR, Axel
: APPLICANT: ZINKE, Holger
: TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
: FILE REFERENCE: 674503-2003
: CURRENT FILING DATE: 1999-06-19
: EARLIER FILING DATE: 1996-06-25
: EARLIER APPLICATION NUMBER: PCT/EP96/02273
: EARLIER FILING DATE: 1995-06-26
: NUMBER OF SEQ ID NOS: 56
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 564
: TYPE: PRT
: ORGANISM: Viscum album
US-08-776-059-35

Query Match
Best Local Similarity 64.9%; Score 387; DB 4; Length 564;
Matches 80; Conservative 12; Mismatches 25; Indels 2; Gaps 1;

QY 1 ARFNPXWRLRRQINSSESSPPNMYMLETISWGHSTQVOQSKDGINFNTQIAISAGN 60
DB 200 ARFNPXWRLRRQINSSESSPPNMYMLETISWGHSTQVOQSKDGINFNTQIAISAGN 259
DB 260 FVTLINVRDVIASLAIMFVCGSRSSSEVRYPVLRVIV--ADDVCSASEPTVRIV 316

RESULT 2
US-08-776-059-39
: Sequence 39, Application US/08776059B
: Patent No. 6271368
: GENERAL INFORMATION:
: APPLICANT: LENTZEN, Hans
: APPLICANT: ECK, Jurgen
: APPLICANT: BAUR, Axel
: APPLICANT: ZINKE, Holger
: TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
: FILE REFERENCE: 674503-2003
: CURRENT APPLICATION NUMBER: US/08/776,059B
```

```

; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Viscum album
US-08-776-059-39

Query Match
Best Local Similarity 71.8%; Score 311; DB 4; Length 235;
Pred. No. 1.9e-32;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 1 ARFNPIKRLRQINSSESPNNMTLETSWGRQSTQVOQSKDGFINTQIRIQISAGN 60
DB 150 ARFNPIKRLRQINSSESPNNMTLETSWGRQSTQVOQSKDGFINTQIRIQISAGN 209
QY 61 FVTVSNVRDVYSSLAIMLFECGSRP 85
DB 210 FVTVSNVRDVYSSLAIMLFECGSRP 234

RESULT 3
US-08-776-059-31
; Sequence 31, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776,059B
; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Viscum album
US-08-776-059-31

Query Match
Best Local Similarity 52.2%; Score 311; DB 4; Length 253;
Pred. No. 2.2e-32;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 1 ARFNPIKRLRQINSSESPNNMTLETSWGRQSTQVOQSKDGFINTQIRIQISAGN 60
DB 168 ARFNPIKRLRQINSSESPNNMTLETSWGRQSTQVOQSKDGFINTQIRIQISAGN 227
QY 61 FVTVSNVRDVYSSLAIMLFECGSRP 85
DB 228 FVTVSNVRDVYSSLAIMLFECGSRP 252

RESULT 4
US-08-378-761A-77
; Sequence 77, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
```

```

; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-761A-77

Query Match
Best Local Similarity 29.7%; Score 177; DB 1; Length 540;
Pred. No. 1.7e-14;
Matches 44; Conservative 18; Mismatches 53; Indels 4; Gaps 1;

QY 1 ARFNPIKRLRQINSSESPNNMTLETSWGRQSTQVOQSKDGFINTQIRIQISAGN 60
DB 178 ARQYIEGEMRTIRYNRSAPDPSPVITLENMGRSLTIQESNOCAGSPITQLORRNOS 237
QY 61 FVTVSNVRDVYSSLAIMLFECGSRPSSLDHPSPLLRSVDAANDVTCTKSEPTVRIY 119
DB 238 KFNVDVSLIPITIALMVKACAPPSSQFS---ILIRVVNFNADVCMDEPIVRIY 292

RESULT 5
US-08-485-286-77
; Sequence 77, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```



```

1 APPLICATION NUMBER: US/08/485,286
2
3 FILING DATE:
4
5 CLASSIFICATION: 435
6
7 PRIOR APPLICATION DATA:
8
9 APPLICATION NUMBER: US 08/378761
10
11 FILING DATE: 26-JAN-1995
12
13 ATTORNEY/AGENT INFORMATION:
14
15 NAME: BORUCKI, ANDREA T
16
17 REGISTRATION NUMBER: 33651
18
19 REFERENCE/DOCKET NUMBER: 38272B
20
21 TELECOMMUNICATION INFORMATION:
22
23 TELEPHONE: (317) 337-4846
24
25 INFORMATION FOR SEQ ID NO: 77:
26
27 SEQUENCE CHARACTERISTICS:
28
29 LENGTH: 540 amino acids
30
31 TYPE: amino acid
32
33 STRANDEDNESS: single
34
35 TOPOLOGY: linear
36
37 MOLECULE TYPE: protein
38
39 US-08-485-286-77

```

RESULT 6
 US-08-356-786-10
 ; Sequence 10, Application US/08356786
 ; Patent No. 5877305
 ; GENERAL INFORMATION:
 ; APPLICANT: Huston, James S.
 ; APPLICANT: Oppermann, Hermann
 ; APPLICANT: Houston, L. L.
 ; APPLICANT: Ring, David B.
 ; TITLE OF INVENTION: BiotSynthetic Binding Protein for Cancer
 ; TITLE OF INVENTION: Marker
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Edmund R. Pletcher, Testa, Hurwitz, & Thibault
 ; STREET: Exchange Place, 53 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,786
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/831,967
 ; FILING DATE: 06-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pletcher, Edmund R.
 ; REGISTRATION NUMBER: 27,829
 ; REFERENCE/DOCKET NUMBER: CRP-053
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 248-7000
 ; TELEFAX: (617) 248-7100
 ;
 ; INFORMATION FOR SEQ ID NO: 10:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 534 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;
US-08-356-786-10

```

Query Match	26.7%;	Score 159;	DB 2;	Length 534;
Best Local Similarity	33.6%;	Pred. No. 3.6e-12;		
Matches 40;	Conservative 23;	Mismatches 50;	Indels 6;	Gaps 2

```

QY      1 ARNPFXKWRRLRRRINSGESPPNNMYLEETFSWMSGOSVVOQSXSDGFNFQILQLTSAN 60
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      182 ARRFYEIGEMRTIRNRNRSAPDPVITLWSMGRLSPAIGESNQGAFAFPDIOLGRNGS 24

```

```

1      RESULT 7
2      US-07-901-707-1
3      ; Sequence 1, Application US/07901707
4      ; Patent No. 5376546
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Bernhard, Susan L.
7      ; APPLICANT: Better, Marc D.
8      ; APPLICANT: Carroll, Steve F.
9      ; APPLICANT: Lane, Julie A.
10     ; TITLE OF INVENTION: Materials Comprising and Methods of
11     ; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
12     ; NUMBER OF SEQUENCES: 57
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
15     ; ADDRESSEE: Bicknell
16     ; STREET: Two First National Plaza, 20 South Clark
17     ; STREET: Street
18     ; CITY: Chicago
19     ; STATE: Illinois
20     ; COUNTRY: USA
21     ; ZIP: 60603
22     ; COMPUTER READABLE FORM:
23     ; MEDIUM TYPE: Floppy disk
24     ; COMPUTER: IBM PC compatible
25     ; OPERATING SYSTEM: PC-DOS/MS-DOS
26     ; SOFTWARE: Patent in Release #1.0, Version #1.25
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/07/901,707
29     ; FILING DATE: 19920619
30     ; CLASSIFICATION: 435
31     ; PRIOR APPLICATION DATA:
32     ; APPLICATION NUMBER: US 07/787,567
33     ; FILING DATE: 04-NOV-1991
34     ; ATTORNEY/AGENT INFORMATION:
35     ; NAME: No. 5376546and, Greta E.
36     ; REGISTRATION NUMBER: 35,302
37     ; REFERENCE/DOCKET NUMBER: 27129/30910
38     ; TELECOMMUNICATION INFORMATION:
39     ; TELEPHONE: (312) 346-5750
40     ; TELEFAX: (312) 984-5750
41     ; TELEFAX: 25-3856
42     ; INFORMATION FOR SEQ ID NO: 1:
43     ; SEQUENCE CHARACTERISTICS:
44     ; LENGTH: 267 amino acids
45     ; TYPE: AMINO ACID
46     ; TOPOLOGY: linear
47     ; MOLECULE TYPE: protein
48     ;
49     ; US-07-901-707-1

```

Query Match	23.2%;	Score 138;	DB 1;	Length 267;
Best Local Similarity	35.6%;	Pred. NO. 7.2e-10;		
Matches	31;	Conservative 18;	Mismatches 38;	Indels 0;
				Gaps 0.

OY 1 ARNP1XWRLRQINSSESPNNMYLETSMGROSTVOYQSKDGFNTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRYNRSADPPSVITLNSWGRSLTAIQSNOGAFASPIQLORRNGS 238
 OY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSYVDVSTILPIIALMWYRCAPPPSS 265

RESULT 8 US-07-988-430-1

; Sequence 1, Application US/07988430
 ; Patent No. 5416202
 ; GENERAL INFORMATION:
 ; APPLICANT: Bernhard, Susan L.
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Lane, Julie A.
 ; APPLICANT: Lei, Shau-Ping
 ; TITLE OF INVENTION: Materials Comprising and Methods of
 ; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
 ; NUMBER OF SEQUENCES: 101
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; STREET: Two First National Plaza, 20 South Clark
 ; STREET: Street
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/988,430
 ; FILING DATE: 19921209
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/901,707
 ; FILING DATE: 19-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5416202and, Greta E.
 ; REGISTRATION NUMBER: 35302
 ; REFERENCE/DOCKET NUMBER: 31133
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 346-5750
 ; TELEFAX: (312) 984-9740
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-988-430-1

Query Match 23.2%; Score 138; DB 1; Length 267;
 Best Local Similarity 35.6%; Pred. No. 7.2e-10;
 Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

OY 1 ARNP1XWRLRQINSSESPNNMYLETSMGROSTVOYQSKDGFNTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRYNRSADPPSVITLNSWGRSLTAIQSNOGAFASPIQLORRNGS 238
 OY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSYVDVSTILPIIALMWYRCAPPPSS 265

RESULT 9 US-08-218-303-16

; Sequence 16, Application US/08218303
 ; Patent No. 5547867
 ; GENERAL INFORMATION:
 ; APPLICANT: Kara, Bhupendra V.
 ; APPLICANT: Hockney, Robert C.
 ; APPLICANT: Filton, John E.
 ; TITLE OF INVENTION: FERMENTATION PROCESS
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cushman, Darcy & Cushman
 ; STREET: 1615 L Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20036-5601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/218,303
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/841,533
 ; FILING DATE: 26-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kokulis, Paul N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: PNK/3893/94908/MJW
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-218-303-16

Query Match 23.2%; Score 138; DB 1; Length 267;
 Best Local Similarity 35.6%; Pred. No. 7.2e-10;
 Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

OY 1 ARNP1XWRLRQINSSESPNNMYLETSMGROSTVOYQSKDGFNTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRYNRSADPPSVITLNSWGRSLTAIQSNOGAFASPIQLORRNGS 238
 OY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSYVDVSTILPIIALMWYRCAPPPSS 265

RESULT 10

; Sequence 1, Application US/08425336
 ; Patent No. 5621083
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Studlika, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 ; PROTEINS
 ; NUMBER OF SEQUENCES: 140
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,336
FILING DATE: 18-APR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Thomas C.
REGISTRATION NUMBER: P-36,989
REFERENCE/DOCKET NUMBER: 31334
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-425-336-1

	Query Match	23.2%	Score 138;	DB 1;	Length 267;
	Best Local Similarity	35.6%;	Pred. No. 7.2e-10;		
	Matches	31; Conservative	18; Mismatches	38; Indels	0; Gaps 0;
Oy	1	A R E N T I X M R L R Q I N S G S S P P N M Y L E L E T S W G S Q S T Q V O G S K D I G F N T Q I R L Q S A C N	60		
		: : : : : : : : : : : : : :			
Db	179	A N F Q I E E G E M R R I R Y N R R S A P D S P V I L E N S W G R L S T A I O S S N O G A F A S P I O L O R N G S	238		
Oy	61	F V T S N V R D V I S L A I M L P E C S G R P F S	87		
	:	: : : : : : : : : : :			
Db	239	K F S V Y D V S I L P T I T A L M V R C A P P S S	265		

```

1      RESULT 11
2      US-08-488-113B-1
3      ; Sequence 1, Application US/08488113B
4      ; Patent No. 5744580
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Better, Marc D.
7      ; APPLICANT: Carroll, Stephen F.
8      ; APPLICANT: Studnicka, Gary M.
9      ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
10     ; TITLE OF INVENTION: Proteins
11     ; NUMBER OF SEQUENCES: 169
12     ; CORRESPONDENCE ADDRESS:
13     ; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
14     ; STREET: 500 West Madison Street, 34th floor
15     ; CITY: Chicago
16     ; STATE: Illinois
17     ; COUNTRY: USA
18     ; ZIP: 60661
19     ; COMPUTER READABLE FORM:
20     ; MEDIUM TYPE: Floppy disk
21     ; COMPUTER: IBM PC compatible
22     ; OPERATING SYSTEM: PC-DOS/MS-DOS
23     ;

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,113B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 110220US07/200-70.P3.C2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-9889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-113B-1

```

	Query Match	Similarity	23.2%	Score 138;	DB 1;	Length 267;
Best Local	Similarity	35.6%	Pred. No. 7.2e-10;			
Matches	31;	Conservative	18;	Mismatches	38;	Indels 0; Gaps 0;
Oy	1	ARFNITXMRRLRQINSGSSPPNNMTMLEETSMGSRQSTOVQOSKDGIFNTQRLQISAGN	60			
Db	179	ARFOYLEGEMRFRIRYNRRSADPSPVITLNSWGRSLTAIOESNOCGAFASPIQLORRNGS	238			
Oy	61	FVTKSNVRDVISLAIMLFECCGRFSS	87			
Db	239	KFSVYDVASILPIIALMVRCAAPSS	265			

RESULT 12
 US-08-477-484B-1
 Sequence 1, Application US/08477484B
 Patent No. 5756899
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Hald & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 City: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

RESULT 13
 US-08-646-360-1
 Sequence 1, Application US/08646360
 Patent No. 5837491
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnicka, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 173
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,360
 FILING DATE: 13-MAY-1996

RESULT 14
 US-08-338-793D-61
 / Sequence 61, Application US/08338793D
 / Patent No. 5840521
 / GENERAL INFORMATION:
 / APPLICANT: Barth, Peter Thomas
 / TITLE OF INVENTION: VECTOR
 / NUMBER OF SEQUENCES: 61
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: CUSHMAN DARBY CUSHMAN
 / ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
 / ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
 / STREET: 1100 New York Avenue, N.W.
 / CITY: Washington
 / STATE: D.C.
 / COUNTRY: U.S.A.
 / ZIP: 20005-3918
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 / COMPUTER: IBM PC/XT/AT Compatibles
 / OPERATING SYSTEM: MS-DOS
 / SOFTWARE: Microsoft Word or ASCII editors
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/338,793D
 / FILING DATE: 08-NO. 5840521-94
 / CLASSIFICATION: 435
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 07/842,081


```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 251:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-765-527-251

Query Match          11.8%; Score 70.5; DB 10; Length 332;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 26; Conservative 23; Mismatches 63; Indels 3; Gaps 2;

OY 1 ARNPFXMRLRQINSGESEPPNMYMLELFSWGRSGNOYQGS-KDGIENFQRLQISAG 59
Db 190 ARRTLENDIRN-NFOQIRPANNNTISLEKNKWKLSFQITSGANWFSEAYELLRAG 247
OY 60 NFTYXSNVLDVYSLSAIMEFECGRPFSSLDHPSPLLIRSVDAANDVCTXSE 114
Db 248 KKYVYAVAVQVVKPIALLKFVDKDKSAACHHHHSRVARRASDEFPSMCAMALDP 302

RESULT 6
US-09-765-527-253
Sequence 253, Application US/09765527
Patent No. US2002000638A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 253:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 253 :
US-09-765-527-253

Query Match 11.7% Score 69.5; DB 10; Length 309;
Best Local Similarity 26.4%; Pred. No. 1.3;
Matches 24; Conservative 19; Mismatches 45; Indels 3; Gaps 2;

QY 1 ARPNDRRLRLQINSGSSPPNNMYMLELETSMGHOSQVOOS-KDGIRNFIQLQISAG 59
||| : : : | - : : : | : ||| : : : | : : : : : : : : : : : : :
DB 190 ARFTFENQIRN-PEQORIRPANTTISLENKMGKLSFQIRTSAGMFSSEAVEERANG 247
: | : | : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 60 NFVTVSNVDVYSSLAIMLFECSSGRPFSSLD 90
: | : | : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 248 KKYVTAVDQYKPKRIALLKFVNDKPKSAALD 278
: | : | : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
US-09-738-626-6445
Sequence 6445, Application US/09738626
Publication NO. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKTO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6445
LENGTH: 477
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6445

Query Match 11.7% Score 69.5; DB 9; Length 477;
Best Local Similarity 35.4%; Pred. No. 2.3;
Matches 17; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

QY 26 MLEETSMGRSTOVQOSKDGIFNFIQLQISAGNVTYSNVDVYSS 73
: : : : | : : : | : : : : | : : : : | : : : : | : : : : | : : : :
DB 241 VLVVDIAGHGHSTMSALK-RIRALDVNPVIVAGNVYTAGDVADIVEA 287
: : : : | : : : | : : : : | : : : : | : : : : | : : : : | : : : :

RESULT 8
US-09-792-793A-34
Sequence 34, Application US/09792793A
Patent No. US20020168370A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Coggin, Philip
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
FILE REFERENCE: 25020-601D

```



```
Query Match          11.0%; Score 65.5; DB 10; Length 937;
Best Local Similarity 29.7%; Pred. No. 17;
Matches 22; Conservative 15; Mismatches 26; Indels 11; Gaps 4;

QY      6  IWMRLRQINSGEESP--PNMYMLELTSWGRSTOVQOSKDGIFN--TQIRLOISAGNFV 62
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      467  LQVNLREIISGTFDPVPNQLFL-----NNLLQAMPSGVSGTLLRLNLSNHF 520

QY      63  T--XSNRVDYISL 74
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      521  SLPVSGVLDQLKSL 534

RESULT 12
US-10-149-819-12
; Sequence 12, Application US/10149819
; Publication No. US20030044913A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dzung Aina M.
; APPLICANT: SHAH, Purvi
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: BUREFORD, Neil
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: EP-0760 PCT
; CURRENT APPLICATION NUMBER: US/10/149,819
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354
; PRIOR FILING DATE: 1999-12-10; 1999-12-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 1110
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: incyte ID No. US20030044913A1 2604493CD1
US-10-149-819-12

Query Match          10.9%; Score 65; DB 9; Length 1110;
Best Local Similarity 25.0%; Pred. No. 25;
Matches 20; Conservative 14; Mismatches 26; Indels 20; Gaps 3;

QY      8  WRLRQINSGEESP-----PNMYMLEL--TSMGRSTOVQOSKDGIFNTOIRLOI 56
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      90  WRSDVPTNTGSKPRVSDAVIDPNNILKTEPEOGTLYSPEOTSLHSESGISGRSSTOM 149

QY      57  S-----AGNFEVTSXNV 67
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      150  NSYSDSGVGEAGSFFHNSQNV 169

RESULT 13
US-09-924-097-15
; Sequence 15, Application US/09924097
; Patent No. US20020156240A1
; GENERAL INFORMATION:
; APPLICANT: TOMONO, Jun
; APPLICANT: NOMURA, Yoshiko
; APPLICANT: SAKAMA, Hiroaki
; APPLICANT: SAKAI, Takeshi
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: ALPHA-AGARASE AND PROCESS FOR PRODUCING THE SAME
; FILE REFERENCE: TOMONO-1
; CURRENT APPLICATION NUMBER: US/09/924,097
; CURRENT FILING DATE: 2001-08-08
```

```
; PRIOR APPLICATION NUMBER: JP00/00966
; PRIOR FILING DATE: 2000-02-21
; PRIOR APPLICATION NUMBER: 11-44890
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: 11-198852
; PRIOR FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of agarase 4-3
US-09-924-097-15

Query Match          10.8%; Score 64.5; DB 9; Length 951;
Best Local Similarity 24.8%; Pred. No. 23;
Matches 27; Conservative 18; Mismatches 51; Indels 13; Gaps 5;

QY      15  NSGESEPPNMYM--ELERSKGRSTOVQOSKDGIFNTOIRLOISAGNFVTSXNVDTYS 72
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      48  NDGQSSPISITVYNGEQALSKVYNAAGDYDVTAV--GGEYNIETLVGTSVAAAPLIEVLV 106

QY      73  SLAIMFECCSGR---PFSSLEHPSPLLRVVDANDVCTXSEPTVRI 118
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      107  D-SNGTFESQGIYIVPLGSMDDFQPLV-----AAHTVTLPAGTSTIRL 148

RESULT 14
US-09-347-064-10
; Sequence 10, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Zinke, Holger
; APPLICANT: Schmidt, Arno
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-10

Query Match          10.7%; Score 64; DB 10; Length 263;
Best Local Similarity 80.0%; Pred. No. 4;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      105  NDVTCXSEPTVRIV 119
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1  DVTGCSASEPTVRIV 15

RESULT 15
US-09-347-064-4
; Sequence 4, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
```

; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-4

Query Match 10.7%; Score 64; DB 10; Length 267;
Best Local Similarity 80.0%; Pred. No. 4.9;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 105 NDVCTXSEPTVRIY 119
:||||:|||||||
Db 1 DDVTCASSEPTVRIY 15

Search completed: March 18, 2003, 08:28:37
Job time : 9.57789 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:55 ; Search time 9.86683 Seconds
(without alignments)
1159.439 Million cell updates/sec

Title: US-09-627-165e-14

Perfect score: 596
Sequence: 1 ARFNP1XWRLRQINSSESS.....VVDANDVCTXSEPTVRIV 119

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	317	53.2	254	2	PD0018
2	178	29.9	576	1	RLCSD
3	177	29.7	564	1	RLCSAG
4	148	24.8	528	1	TRISA
5	138	23.2	528	2	S32431
6	138	23.2	562	2	S16022
7	135	22.7	527	2	S32430
8	121.5	20.4	570	2	S62627
9	99.5	16.7	106	2	B39761
10	93.5	15.7	251	2	C39761
11	92	15.4	286	1	R8PUGC
12	84	14.1	294	2	S28421
13	79	13.3	261	2	JE0401
14	78	13.1	247	2	JC5032
15	76.5	12.8	289	2	JC5606
16	76	12.8	247	2	JU0393
17	76	12.8	754	2	S48020
18	74	12.4	277	2	S24494
19	74	12.4	313	2	S17757
20	74	12.4	730	2	T43943
21	72.5	12.2	289	1	RLRZT
22	71	11.9	797	2	A87247
23	70	11.7	730	2	T44066
24	69.5	11.7	1054	2	A61221
25	68.5	11.5	268	2	S30135
26	68.5	11.5	316	2	JT0753
27	68	11.4	467	2	S59430
28	68	11.4	1957	2	A45627
29	67.5	11.3	250	2	JN0108

30	67.5	11.3	278	2	S23519	beta-luffin - smoo
31	67	11.2	572	1	B43675	64k protein - infe
32	66.5	11.2	245	2	JC4840	rRNA N-glycosidase
33	66.5	11.2	2257	2	T09538	acetyl-CoA carboxy
34	66	11.1	682	2	B84415	translation elonga
35	66	11.1	728	1	S07558	flagellar p-ring p
36	65.5	11.0	385	2	A84970	probable glutamate
37	65.5	11.0	462	2	C90724	acyl carrier prote
38	65.5	11.0	462	2	C85575	DNA translocase st
39	65.5	11.0	38	2	A69254	insulin receptor s
40	65.5	11.0	467	2	C69999	lambda 3 protein -
41	65.5	11.0	702	2	S30185	lambda 3 protein -
42	65.5	11.0	1231	1	MMXR31	probable cell divi
43	65.5	11.0	1267	1	MMXR33	telomerase-associ
44	65.5	11.0	1305	2	AB0168	
45	65.5	11.0	2629	2	T32735	

ALIGNMENTS

RESULT 1

PD0018 mistletoe lectin I A chain - Viscum album (fragment)

C:Species: Viscum album

C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 26-Aug-1999

C:Accession: PD0018

R:Eschenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, B.

Biochem. Biophys. Res. Commun. 247, 367-372, 1998

A:Title: Primary structure and molecular modelling of mistletoe lectin I from Viscum

A:Reference number: PD0018; MUID:98308123; PMID:9642133

A:Accession: PD0018

A:Molecule type: protein

A:Residues: 1-254 <ESC>

A:Superfamily: ricin; rRNA N-glycosidase homology

F:7-246/Domain: rRNA N-glycosidase homology <RMG>

Query Match Best Local Similarity 71.6% ; Score 317; DB 2; Length 254; Pred. No. 3.9e-28; Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 1 ARFNP1XWRLRQINSSESSPPNKMTELETSKGRQSTOVQOQSDGIFNPQIRLOISAGN 60

DB 167 ARFNP1XWRLRQINSSESSFLPDVYMELRSTWGOQSTOVQHSYDGVFNNPDIRLALPPGN 226

QY 61 FVTLXSWRDVTSIAIMLFECGSRPSS 88

DB 227 FVTLXSWRDVTSIAIMLFECGSRPSS 254

RESULT 2

RLCSD ricin D precursor - castor bean

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis

A:Reference number: A24041; MUID:86067214; PMID:2999712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin g

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TR>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R.Lamb, F.I.: Roberts, L.M.; Lord, J.M.
 Eur. J. Biochem. 148, 265-270, 1985
 A:Title: Nucleotide sequence of cloned cDNA coding for preprotricin.
 A:Reference number: A24614; MUID:85179479; PMID:3838723
 A:Accession: A24614
 A:Molecule type: mRNA
 A:Residues: 12-75,'D','77-550','R','552-576 <LAM>
 A:Cross-references: GB:X02388; MID:g21077; PIDN:CAA26230.1; PID:g21078
 R.Yoshitake, S.; Funatsu, G.; Funatsu, M.
 Agric. Biol. Chem. 42, 1267-1274, 1978
 A:Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile of
 A:Reference number: A03372
 A:Accession: A03372
 A:Molecule type: Protein
 A:Residues: 36-97,'Q','99-109','S','111-269','D','272-283','L','285-288,290-302 <YOS>
 A:Note: This paper cites the others in the series providing experimental details for the
 R:Ataki, T.; Funatsu, G.
 FEBS Lett. 191, 121-124, 1985
 A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan
 A:Reference number: A24010
 A:Accession: A24010
 A:Molecule type: Protein
 A:Residues: 315-383,'PS','386-576 <ARA>
 R:Funatsu, G.; Kimura, M.; Funatsu, M.
 Agric. Biol. Chem. 43, 2221-2224, 1979
 A:Title: Primary structure of Ala chain of ricin D.
 A:Reference number: A03374
 A:Accession: A03374
 A:Molecule type: Protein
 A:Residues: 315-335,'N','337-342','NH','345-362,364-383,'PS','386-399','T','401','D','403','E','40
 527','E','529-564','W','566','H','567-570','LT','573-574','P' <YUN>
 A:Note: This paper, one of a series, summarizes the experimental details for the determi
 R:Ready, M.P.; Kim, Y.; Robertus, J.D.
 Proteins 10, 270-278, 1991
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism o
 A:Reference number: A48237; MUID:91352006; PMID:1881883
 A:Contents: annotation: active site
 R:Rutember, E.; Robertus, J.D.
 Proteins 10, 260-269, 1991
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.
 A:Reference number: A48236; MUID:91352005; PMID:1881882
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms
 R:Katin, B.-J.; Collins, E.J.; Robertus, J.D.
 Proteins 10, 251-259, 1991
 A:Title: Structure of ricin A-chain at 2.5 angstroms.
 A:Reference number: A48239; MUID:91352004; PMID:1881881
 A:Contents: annotation: X-ray crystallography, 2.5 angstroms
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which
 C:Comment: The A chain inhibits protein synthesis; it inactivates the 60S ribosomal subu
 C:Comment: This protein is cytotoxic and very poisonous to animals.
 C:Superfamily: ricin: RNA N-glycosidase homology
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-302/Product: ricin D chain A #status experimental <ACH>
 F:46-293/Domain: RNA N-glycosidase homology <RNG>
 F:315-576/Product: ricin D chain B #status experimental <BCH>
 F:331-373,374-414,417-455,462-497,501-540,543-576/Region: 40-residue repeats
 F:45,409,449/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:115,158,243,244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:212/Active site: Glu #status experimental
 F:215/Active site: Arg #status predicted
 F:294-318,334-353,377-394,465-478,504-521/Disulfide bonds: #status experimental
 F:336,349,360/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status experimental
 F:548,569/Binding site: N-acetylgalactosamine (Asp, Asn) #status experimental

Query Match 29.98; Score 178; DB 1; Length 576;
 Best Local Similarity 37.08; Pred. No. 6e-12;
 Matches 44; Conservative 19; Mismatches 52; Indels 4; Gaps 1;

QY 1 ARFNPXRLRQINSSESPNNYMLEFTSMGROSTOVQOSKDGIFNTQRLQISAGN 60
 DB 214 ARFOYIGEMRTIRRYNRSAAPDPSPVITLNSWGRSLTAIOESNOGAFASPILOQRNGS 273

QY 61 FVTXSNVADVSSLAIMLEFSGRPFSSLDHPSPILRLSVDAANDVCTXSEPTVRIY 119
 DB 274 KFSYVDVSLIPITAILMWYRCAPPPSSQFS-----LLIRPVYPMFNADVCMDEPIVRIY 328

RESULT 3
 RLCSAG
 agglutinin precursor - castor bean
 N:Contains: RNA N-glycosidase (EC 3.2.2.22)
 C:Species: Ricinus communis (castor bean)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: A24261; A24210
 R:Roberts, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.
 J. Biol. Chem. 260, 15682-15686, 1985
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.
 A:Reference number: A24261; MUID:86059449; PMID:299130
 A:Accession: A24261
 A:Molecule type: mRNA
 A:Residues: 1-564 <ROB>
 A:Cross-references: GB:M12089; NID:g169700; PIDN:AAA3869.1; PID:g169701
 R:Ataki, T.; Yoshioaka, Y.; Funatsu, G.
 Biochim. Biophys. Acta 872, 277-285, 1986
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis aggl
 A:Reference number: A24210
 A:Accession: A24210
 A:Molecule type: Protein
 A:Residues: 303-325,'F','327-330','T','332-361','D','363-373','G','375-403','T','405-551','V','5
 C:Comment: This protein has strong agglutinating activity and weak cytotoxicity compa
 C:Superfamily: ricin: RNA N-glycosidase homology
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; s
 F:25-290/Product: signal sequence #status predicted <SIG>
 F:35-281/Domain: agglutinin chain A #status predicted <ACH>
 F:303-564/Product: agglutinin chain B #status experimental <BCH>
 F:319-361,362-402,405-443,450-485,489-528,531-564/Region: 40-residue repeats
 F:34,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:104,147,231,232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:200,203/Active site: Glu, Arg #status predicted
 F:282-306,322-341,365-382,453-466,492-509/Disulfide bonds: #status predicted
 F:324,337,348/Binding site: N-acetylgalactosamine (Asp, Glu, Asn) #status predicted
 F:397,437/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:536,557/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 29.78; Score 177; DB 1; Length 564;
 Best Local Similarity 37.08; Pred. No. 7.5e-12;
 Matches 44; Conservative 18; Mismatches 53; Indels 4; Gaps 1;

QY 1 ARFNPXRLRQINSSESPNNYMLEFTSMGROSTOVQOSKDGIFNTQRLQISAGN 60
 DB 202 ARFOYIGEMRTIRRYNRSAAPDPSPVITLNSWGRSLTAIOESNOGAFASPILOQRNGS 261

QY 61 FVTXSNVADVSSLAIMLEFSGRPFSSLDHPSPILRLSVDAANDVCTXSEPTVRIY 119
 DB 262 KFNVDVSLIPITAILMWYRCAPPPSSQFS-----LLIRPVYPMFNADVCMDEPIVRIY 316

RESULT 4
 TRLSA
 abrin-a precursor - Indian licorice (fragment)
 N:Contains: RNA N-glycosidase (EC 3.2.2.22)
 C:Species: Abrus precatorius (Indian licorice)
 C:Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999
 C:Accession: S32429; J0202; A39761; J01398; S14472; S24133; S74110; S74111
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
 J. Mol. Biol. 229, 263-267, 1993
 A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.
 A:Reference number: S32429; MUID:93132798; PMID:8421313
 A:Accession: S32429
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 'E',2-528 <HUN>
 A:Cross-references: GB:M98344; NID:g166294; PIDN:AAA32624.1; PID:g166295

A:Note: the coding region for the sequence shown is preceded by an ATG codon

A:Note: residues 1-8 were derived from the synthesized primer

R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Tanaka, M.

Agri. Biol. Chem. 52, 1095-1097, 1988

A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from

A:Reference number: JT0202

A:Accession: JT0202

A:Molecule type: protein

A:Residues: 1-201,203-251 <FUN>

A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have

R:Evensen, G.; Mathiesen, A.; Sundan, A.

J. Biol. Chem. 266, 6848-6852, 1991

A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.

A:Reference number: A39761; MUID:91201329; PMID:2016300

A:Accession: A39761

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 'E',2-251 <EVE>

A:Cross-references: GB:X54872

A:Note: residues 1-8 were derived from the synthesized primer

R:Kimura, M.; Sumizawa, T.; Funatsu, G.

Bioosci. Biotechnol. Biochem. 57, 166-169, 1993

A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic

A:Reference number: JC1398; MUID:93169023; PMID:7763422

A:Contents: seeds

A:Accession: JC1398

A:Molecule type: protein

A:Residues: 261-347,'T',349-351,'A',353-357,'L',359-528 <KIM>

A:Experimental source: seed

R:Evensen, G.; Mathiesen, A.; Sundan, A.

submitted to the EMBL Data Library, October 1990

A:Description: Direct molecular cloning of two distinct abrin A-chains.

A:Reference number: S14471

A:Accession: S14472

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'ME',2-251 <EV2>

A:Cross-references: EMBL:X54873; NID:q16090; PIDN:CAA3655.1; PID:q16091

R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.

PNAS Lett. 309, 115-118, 1992

A:Title: The complete primary structure of abrin-a B chain.

A:Reference number: S24133; MUID:92371656; PMID:1505674

A:Accession: S24133

A:Molecule type: protein

A:Residues: 262-297,'Y',299-426,'L',428-466,'P',468-482,'L',484-528 <CH>

R:Lin, S.H.; Chow, Y.P.; Chen, Y.L.; Law, Y.C.; Chen, J.K.; Lin, J.Y.

Eur. J. Biochem. 240, 564-569, 1996

A:Title: Probing the domain structure of abrin-a by tryptic digestion.

A:Reference number: S74110; MUID:97008945; PMID:8856055

A:Accession: S74110

A:Molecule type: protein

A:Residues: 89-108;154-172 <LIN>

A:Experimental source: seed

A:Accession: S74111

A:Molecule type: protein

A:Residues: 262-276,'X',278-280;329-348;369-388;399-418 <LIM>

A:Experimental source: seed

C:Comment: Abria-a is more toxic than ricin. The toxin consists of an A chain, which contains receptors on the cell surface. The A and B chains are linked by a single disulfide

C:Superfamily: ricin; rRNA N-glycosidase homolog

C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyutamic acid

F:1-251/Product: abrin-a chain A #status experimental <ACH>

F:1-246/Domain: rRNA N-glycosidase homolog <RNG>

F:261-528/Product: abrin a chain B #status experimental <BOH>

F:283-325;326-366;369-407;414-449;453-492;495-528/Region: 40-residue repeats

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted

F:247-269;286-305;329-346;417-430;456-473/Disulfide bonds: #status predicted

F:361,401/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 24.8% Score 148; DB 1; Length 528;
Best Local Similarity 33.3%; Pred. No. 1,3e-08;
Matches 43; Conservative 23; Mismatches 37; Indels 26; Gaps 5;

OY 1 ARFPDIXWRLRRQINSGSSPPNNMYMLEETSMGROSTOVQOOSKDCIFNTQIRLOISAGN 60
||| : || : | : | : | : | : | : | : | : | : | : | : | : | :
Db 166 ARFRISRVRVSIQTGTAFQPDAMITSLENNMDLISRGVOESYODTFEPNQ----- 216

OY 61 FVTYSNVARD-----VISSLAIMLFECDSGRPFSSLDHPSPLLRSLRVDAANDVC 109
|| : || : | : | : | : | : | : | : | : | : | : | : | : | :
Db 217 -VTLTINNPEPVIDLSHPTVAVALMLFVCN--PENA--NOSPILRSIVEESK-ICS 270
OY 110 TXSEPTVRI 118
:|||||:
Db 271 SRYEPTVRI 279

RESULT 5
S32431
abrin-d precursor - Indian licoice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licoice)
C:date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
C:Accession: S32431; S34408
J:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J:MOL. Biol. 229, 263-267, 1993
A:title: Primary structure of three distinct isoforms determined by cDNA sequencing
A:Reference number: S32429; MUID:93132798; PMID:8422133
A:Accession: S32431
A:molecule type: mRNA
A:Residues: 1-528 <HUN>
A:CROSS-references: GB:M98346
R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
submitted to the EMBL Data Library, March 1993
A:Reference number: S34408
A:Accession: S34408
A:molecule type: mRNA
A:Residues: 1-169, 'C', 171-320, 'U', 322-528 <HUN>
A:CROSS-references: GB:M98346
C:Comment: abrin consists of an A chain, which inhibits protein synthesis by inactive
The A and B chains are linked by a single disulfide bond, which is essential for to
C:Keywords: disulfide bond; duplication; glycoprotein; glycosylase; hydrolase; lectin
F:1-251/Product: abrin d chain A #status predicted <CH>
F:7-246/Domain: rRNA N-glycosidase homology <NG>
F:7-246/Product: abrin d chain B #status predicted <BCH>
F:261-528/Product: abrin d chain B #status predicted
F:283-320,326-366,369-407,411-430,456-473/Disulfide bonds: #status predicted
F:1,Modified site: pyrroldione carboxylic acid (Glu) #status predicted
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted
F:200,253,361,401,402/Binding site: carbonylate (Asn) (covalent) #status predicted
F:247-269,286-305,329-346,411-430,456-473/Disulfide bonds: #status predicted
F:288,312/Binding site: N-acetylglucosamine (ASP, Asn) #status predicted
F:500,521/Binding site: N-acetylglucosamine (ASP, Asn) #status predicted

Query Match 23.2% Score 138; DB 2; Length 528;
Best Local Similarity 36.1%; Pred. NO. 1.8e-07;
Matches 43; Conservative 20; Mismatches 50; Indels 6; Gaps 4;

OY 1 ARFNPIXWLRQINSGSSPPNMYLEETSMGROSTOVQOOSKDGI-FNTQIRLOISAG 59
|| : || : | : | : | : | : | : | : | : | : | : | : | : | :
Db 166 ARFRYSNRVGVSIFRTGARQPDPAMLSLENNMDNLISGVGOVSODTFEPNNVLISSTRQ 225

OY 60 NFPTYXNRVDVISLAIMLFECGSRPFSSLDHPSPLLRSLRVDAANDVCXTXSPTVRI 118
|| : || : | : | : | : | : | : | : | : | : | : | : | : | :
Db 226 PAVVDLSHSHTPAVAILMLFVCN--PENNA--NOSPILRSIVEESK-ICSSRYEPTVRI 279

RESULT 6
S16022
abrin-c precursor - Indian licoice
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licoice)

```

F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
F:287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match          22.7%; Score 135; DB 2; Length 527;
Best Local Similarity 33.3%; Pred. No. 3,9e-07;
Matches 40; Conservative 24; Mismatches 48; Indels 8; Gaps 5;

QY      1  ARENPIDXRRLROINSGESSPPNMVMELETMSGRSTOVQOSKCIFFQTQRLDISAGN 60
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      165 ARRFRIYRQVGSIRFTYTAQPDAAMISLENNMNDISGGVQSDPFPNAVTLR-SVNN 223
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      224 QPVIDSLFTHQSAVALMLMFVCN--PPNA--NQSPLLIRLIVKSK-ICSSRYEPTVRI 278
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
S62627
agglutinin I precursor - European elder
C:Species: Sambucus nigra (European elder)
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
C:Accession: S62627; S62619
R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Eur. J. Biochem. 225, 128-137, 1996
A:Title: The NeuAc(alpha-2,6)-Gal/gal1NAc-binding lectin from elderberry (Sambucus nigra)
A:Reference number: S62619; MUID:96202926; PMID:8631319
A:Accession: S62627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-570 <V>A>
A:Cross-references: EMBL:U27122; NID:g1141772; PID:MAC49158.1; PID:g1141773
A:Accession: S62619
A:Status: preliminary
A:Molecule type: protein
A:Residues: 29-39;309-319 <V>A>
F:Superfamily: ricin; RNA N-glycosidase homology <RNG>
F:37-283/Domain: RNA N-glycosidase homology <RNG>

Query Match          20.4%; Score 121.5; DB 2; Length 570;
Best Local Similarity 25.2%; Pred. No. 1.4e-05;
Matches 32; Conservative 31; Mismatches 53; Indels 11; Gaps 4;

QY      1  ARENPIDXRRLROINSGESSPPNMVMELETMSGRSTOVQOSK-DGIFNTQRLQISAG 59
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      201 ARRFRIELRIRKISTIDASEPTPDLMLSMENKMSMSSEIQDQPGGIFAGVOLRDERN 260
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      60 NFVYKSNVDV--ISLAIMEFGSGRPFSLDHPSPILLRSVY-----DAANDVCTX 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      261 NSIEVTNRRRLFEELYIAVLVYGA--PVTSSYSNNAIDMQIKMVPFGGEYKCVS 318
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      112 SEPTVRI 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      319 VEVTIRRI 325

RESULT 9
B39761
abrin (clone 3-7) precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 05-Jun-1998
C:Accession: B39761
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329; PMID:2016500
A:Accession: B39761
A:Molecule type: DNA
A:Residues: 1-106 <E>E>
A:Cross-references: GB:X54873
F:Superfamily: ricin; RNA N-glycosidase homology
F:Keywords: duplication; rRNA N-glycosidase; hybrid

```

RESULT 9
B39761
abrin (clone 3.7) precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 05-Jun-1998
C:Accession: B39761
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
M:Title: Direct molecular cloning and expression of two distinct abrin A-chains
A:Reference number: A39761; MUID:91201329; PMID:2016300
A:Accession: B39761
A:Molecule type: DNA
A:Residues: 1-106 <EYE>
A:Cross-references: GB:X54873
C:Superfamily: ricin; rRNA N-glycosidase homology
#Keywords: duplication; glycosidase; hydrolase; lectin; toxin

F.1-106/Product: abrin chain A (fragment) #status predicted <ACH>
F.1-101/Domain: rRNA N-glycosidase homology (fragment) <RNG>
F.19,22/Active site: Glu, Arg #status predicted

A:Molecule type: DNA
A:Residues: 1-294 <KAT>
A:Cross-references: EMBL:D10600; NID:g218010; PIDN:BAA01451.1; PID:g218011
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:Keywords: glycosidase; hydrolase
F:30-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 14.1%; Score 84; DB 2; Length 294;
Best Local Similarity 26.6%; Pred. No. 0.11;
Matches 21; Conservative 19; Mismatches 37; Indels 2; Gaps 1;

OY 1 ARFNPIXRRLRQINSSESPNNMYLETSMGROSTOVQOOSKDGIFNTQIRLOISA 60
DB 201 ARFKYIENQVKTNN--RAFPNNAVLNLESMGKISTAIHNAKNGALTSPLELNANGS 258
OY 61 EYTXSNRVDVTSILMIF 79
DB 259 KWIVLRVDIDPVDLLKAT 277

RESULT 13
JE0401
antiviral protein - Virginian pokeweed
C:Species: Phytolacca americana (Virginian pokeweed)
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 07-May-1999
C:Accession: JE0401
R:Kung, S.S.; Kimura, M.; Funatsu, G.
Agric. Biol. Chem. 54, 3301-3318, 1990
A:Title: The complete amino acid sequence of antiviral protein from the seeds of pokeweed
A:Reference number: JE0401; MUID:91242096; PMID:1368643
A:Accession: JE0401
A:Molecule type: protein
A:Residues: 1-261 <KUN>
A:Experimental source: seed
C:Comment: This protein prevents the replication of a number of plant viruses, and inhibits
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:Keywords: disulfide bond; glycoprotein
F:6-254/Domain: rRNA N-glycosidase homology <RNG>
F:10/44,255/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:34-258,84-105/Disulfide bonds: #status experimental

Query Match 13.3%; Score 79; DB 2; Length 261;
Best Local Similarity 33.9%; Pred. No. 0.34;
Matches 20; Conservative 13; Mismatches 25; Indels 2; Gaps 1;

OY 1 ARFNPIXRRLRQINSSESPNNMYLETSMGROSTOVQOOSKDGIFNTQIRLOISA 59
DB 177 ARFYIENQVKTNNRDFS--PNDKVLDEENMGKISTAIHNSKNGALPPLDLKNDG 233

RESULT 14
JC5032
karasurin-B - Trichosanthes kirilowii var. japonica
C:Species: Trichosanthes kirilowii var. japonica
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 23-May-1997
C:Accession: JC5032
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
Biol. Pharm. Bull. 19, 1485-1489, 1996
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-A
A:Reference number: JC5032; MUID:97108848; PMID:8951169
A:Accession: JC5032
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-247 <KON>
C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 13.1%; Score 78; DB 2; Length 247;
Best Local Similarity 28.0%; Pred. No. 0.41;
Matches 23; Conservative 23; Mismatches 30; Indels 6; Gaps 3;

OY 1 ARFNPIXRRLRQINSSESPNNMYLETSMGROSTOVQ--QSKDGIFFNTQIRLOISA 58

DB 162 ARFYIEQIQIGKRD--KTFPLSLATISLENSALSQIOIASINNGOFETPVVLINAO 219
OY 59 GNFVYTXSNVRD--VTSILMIF 78
DB 220 NORVITITNVAGVVTSMIALLL 241

RESULT 15
JC5606
karasurin C - Trichosanthes kirilowii var. japonica
N:Contains: karasurin A
C:Species: Trichosanthes kirilowii var. japonica
C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 19-Jul-2002
C:Accession: JC5606; JC5033
R:Mizukami, H.; Iida, K.; Kondo, T.; Ogihara, Y.
Biol. Pharm. Bull. 20, 711-713, 1997
A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating protein
A:Reference number: JC5606; MUID:97356562; PMID:9212998
A:Accession: JC5606
A:Molecule type: DNA
A:Residues: 1-289 <MTZ>
A:Cross-references: DDBJ:AB000666; NID:g2329830; PIDN:BAA21786.1; PID:g2329831
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
Biol. Pharm. Bull. 19, 1485-1489, 1996
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-A
A:Reference number: JC5032; MUID:97108848; PMID:8951169
A:Accession: JC5033
A:Status: preliminary
A:Molecule type: protein
A:Residues: 22-270 <KON>
C:Comment: This protein is a ribosome-inactivating protein and exhibits cytotoxic, abortifacient, and
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:22-270/Product: karasurin C #status predicted <MAC>
F:24-270/Product: karasurin A #status predicted <MAC>
F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 12.8%; Score 76.5; DB 2; Length 289;
Best Local Similarity 25.3%; Pred. No. 0.74;
Matches 25; Conservative 28; Mismatches 37; Indels 9; Gaps 4;

OY 1 ARFNPIXRRLRQINSSESPNNMYLETSMGROSTOVQ--QSKDGIFFNTQIRLOISA 58
DB 185 ARFKYIEQIQIGKRD--KTFPLSLATISLENSALSQIOIASINNGOFETPVVLINAO 242
OY 59 GNFVYTXSNVRD--VTSILMIFECGSGRPSSLDHPSP 95
DB 243 NORVITITNVAGVVTSMIALLL---NRNMMAIDDDVPW 278

Search completed: March 18, 2003, 08:27:30
Job time : 15.8668 secs

FT VARIANT 231 231 T -> S (IN M.L.A.).
 FT VARIANT 235 235 D -> S (IN M.L.A.).
 SO SEQUENCE 254 AA; 28478 MW; 53BAPF8D3E0FF67 CRC64;

Query Match 53 2%; Score 317; DB 1; Length 254;
 Best Local Similarity 71.6%; Pred. No. 3.7e-29;
 Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

DB 1 ARPNFRLRQINSSESPPMVYMLETSGRSTQVOQSKDIPNTQIRLQISAGN 60
 167 ARFNILMRKYROYINSAGSFLPDVYMLETSGMGOSTQOVHSTDGVEFNPRILAIIPGN 226
 OY 61 FVTSNVRDVTSIAIMLFECSSGRPSS 88
 DB 227 FVTLNVRDVTSIAIMLFVCGRSPSS 254

RESULT 2
 RIC1_RICCO STANDARD; PRT; 576 AA.
 AC P02879; P02880;

DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ricin precursor [contains: Ricin A chain (RNA N-glycosidase)
 DE (EC 3.2.2.22); Ricin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Ericosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_Taxid=3988;

RI MEDLINE=86067214; PubMed=2999712;
 RA Hailing K.C., Hailing A.C., Murray E.E., Ladin B.F., Houston L.L.,
 RA Weaver R.F.;
 RT "Genomic cloning and characterization of a ricin gene from Ricinus
 RT communis";
 RL Nucleic Acids Res. 13:8019-8033(1985).

RI MEDLINE=92163016; PubMed=13711405;
 RA Tregear J.W., Roberts L.M.;
 RT "The lectin gene family of Ricinus communis: cloning of a functional
 RT ricin gene and three lectin pseudogenes.";
 RL Plant Mol. Biol. 18:515-525(1992).
 RN [3]
 RP SEQUENCE OF 12-576 FROM N.A.;
 RP MEDLINE=85179479; PubMed=3838723;
 RA Lamb A., Roberts L.M., Lord J.M.;
 RT "Nucleotide sequence of cloned cDNA coding for preproricin";
 RL Eur. J. Biochem. 148:265-270(1985).

RI MEDLINE=90344223; PubMed=1368517;
 RA Yoshitake S., Funatsu G., Funatsu M.;
 RT "Isolation and sequences of peptic peptides, and the complete
 RT sequence of the chain of ricin-D.";
 RL Agric. Biol. Chem. 42:1267-1274(1978).

RI MEDLINE=21480122; PubMed=11595634;
 RA Olsnes S., Kozlov J.V.;
 RP REVIEW
 RT "Structural analyses of sugar chains from ricin A-chain variant";
 RL Agric. Biol. Chem. 54:157-162(1990).

RI MEDLINE=90344223; PubMed=1368517;
 RA Yoshitake S., Funatsu G., Funatsu M.;
 RT "Isolation and sequences of peptic peptides, and the complete
 RT sequence of the chain of ricin-D.";
 RL Agric. Biol. Chem. 42:1267-1274(1978).

RT "Ricin";
 RL Toxicol. 39:1723-1728(2001).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RA MEDLINE=87165983; PubMed=3558397;
 RA Monfort W., Villafraña J.E., Monzinger A.F., Ernst S.R., Katzin B.,
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;
 RT "The three-dimensional structure of ricin at 2.8 A.";
 RL J. Biol. Chem. 262:5398-5403(1987).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
 RX MEDLINE=91352004; PubMed=1881881;
 RA Katzin B.J., Collins E.J., Robertus J.D.;
 RT "Structure of ricin A-chain at 2.5 A.";
 RL Proteins 10:251-259(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
 RX MEDLINE=91352005; PubMed=1881882;
 RA Rutenber E., Robertus J.D.;
 RT "Structure of ricin B-chain at 2.5-A resolution.";
 RL Proteins 10:260-269(1991).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
 RX MEDLINE=95082010; PubMed=7990130;
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,
 RA Pauprit R.A.;
 RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";
 RL J. Mol. Biol. 244:410-422(1994).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
 RX MEDLINE=96574422; PubMed=8760513;
 RA Day P.J., Ernst S.R., Frankel A.E., Monzinger A.F., Pascal J.M.,
 RA Molina-Svinth M.C., Robertus J.D.;
 RT "Structure and activity of an active site substitution of ricin A
 RT chain";
 RL Biochemistry 35:11098-11103(1996).
 RN [13]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
 RX MEDLINE=97240820; PubMed=9066280;
 RA Yan X., Hollis T., Svinth M., Day P., Monzinger A.F., Milne G.W.,
 RA Robertus J.D.;
 RT "Structure-based identification of a ricin inhibitor";
 RL J. Mol. Biol. 266:1043-1049(1997).
 RN [14]
 RP MOTAGENESIS.
 RX MEDLINE=93165632; PubMed=1287657;
 RA Kin Y., Robertus J.D.;
 RT "Analysis of several key active site residues of ricin A chain by
 RT mutagenesis and X-ray crystallography";
 RL Protein Eng. 5:775-779(1992).
 CC -!- FUNCTION: Ricin is highly toxic to animal cells and to a less
 CC extent to plant cells. The A chain is responsible for inhibiting
 CC protein synthesis through the catalytic inactivation of 60S
 CC ribosomal subunits. It acts as a glycosylase that removes a
 CC specific adenine residue from an exposed loop of 28S ribosomal
 CC RNA. As this loop is involved in the binding of elongation
 CC factors, the modified ribosomes are unable to support protein
 CC synthesis. The A chain can inactivate a few thousand ribosomes
 CC per minute, thus inactivating them faster than the cell can make
 CC new ones. A single A-chain molecule can therefore kill an animal
 CC cell. The B chain binds to cell receptors and facilitates the
 CC entry into the cell of the A chain. B chains are also responsible
 CC for cell agglutination (lectin activity). It binds to beta-D-
 CC galactopyranoside moieties.
 CC -!- CATALYTIC ACTIVITY: Endohydrolisis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -!- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

```

CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC -1- CAUTION: REF. 4 AND REF. 5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS
CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; X03179; CAA26939.1; -
DR EMBL; X52908; CAA37095.1; -
DR EMBL; X02388; CAA26230.1; -
DR EMBL; A12892; CAA01058.1; -
DR PIR; A24041; RUCSD.
DR PDB; 2AAT; 31-JAN-94.
DR PDB; 1APG; 31-JAN-94.
DR PDB; 1FMP; 31-OCT-93.
DR PDB; 1IFS; 14-JAN-98.
DR PDB; 1IFW; 14-JAN-98.
DR PDB; 1IFJ; 14-JAN-98.
DR PDB; 1RTC; 31-OCT-93.
DR PDB; 1OBS; 16-JUN-97.
DR PDB; 1OBT; 16-JUN-97.
DR PDB; 1BR5; 02-SEP-98.
DR PDB; 1BR6; 02-SEP-98.
DR GlycoSuiteDB; P02879; -.
DR InterPro; IPR001574; RIP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00161; RIP; 1
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Signal; 3D-structure.
FT CHAIN 1 35
FT SIGNAL 1 35
FT PEPTIDE 303 314
FT CHAIN 315 576
FT DOMAIN 321 448
FT REPEAT 331 373
FT REPEAT 374 414
FT REPEAT 417 449
FT REPEAT 462 497
FT REPEAT 501 540
FT REPEAT 543 570
FT ACT_SITE 212 212
FT DISULFID 294 318
FT DISULFID 334 353
FT DISULFID 377 394
FT DISULFID 465 478
FT DISULFID 504 521
FT CARBOHYD 45 45
FT CARBOHYD 271 271
FT CARBOHYD 409 409
FT CARBOHYD 449 449
FT CONFLICT 76 76
FT CONFLICT 551 551
FT STRAND 43 47
FT TURN 49 50
FT TURN 53 57
FT HELIX 53 67
FT STRAND 73 74
FT TURN 75 76
FT TURN 77 79
FT STRAND 88 90
FT TURN 88 90

```

[illegible]

Query Match	29.7%;	Score 177;	DB 1;	Length 564;
Best Local Similarity	37.0%;	Pred. No. 1.4e-12;		
Matches	44;	Conservative	18;	Mismatches 53;
			Indels	4;
			Gaps	1

RESULT 4	
ABRA_ABRPR	STANDARD;
ID ABRA_ABRPR	PRT; 528 AA.

DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)]
DE (RC 3.2.2.22): Abrin-a B chain)
OS Abrus precatorius (Indian licorice) ('Crab's eye').
OC Embryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae
OC eucotsids I; Finales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93132798; PubMed=8421313;
RX Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
RA "Primary structure of three distinct isoabins determined by cDNA
RT sequencing. Conservation and significance.";
RL J. Mol. Biol. 229:263-267(1993).
RN [2]
RN SEQUENCE OF 1-251.
RC TISSUE=Seed:
RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic

PT		protein from the seeds of Abrus precatorius.";
RL		Agric. Biol. Chem. 52:1095-1097(1988).
RN	[3]	SEQUENCE OF 1-251 FROM N.A.
RP		TISSUE=Leaf;
RC		MEDLINE=91201329; PubMed=2016300;
RX		Evensen G., Mathiesen A., Sundan A.;
RA		"Direct molecular cloning and expression of two distinct abrin
RT		A-chains";
RL	J. Biol. Chem.	266:6848-6852(1991).
RN	[4]	
RP		SEQUENCE OF 262-528.
RC		MEDLINE=92371656; PubMed=1505674;
RX		Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
RA		"The complete primary structure of abrin-a B chain.";
RT		FEBS Lett. 309:115-118(1992).
RL		
RN	[5]	X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
RP		MEDLINE=95333188; PubMed=7608980;
RX		Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RA		"Crystal structure of abrin-a at 2.14 A.";
RT	J. Mol. Biol.	250:354-367(1995).
RL		-I- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC		SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC		SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC		ABRIN-A IS MORE TOXIC THAN RICIN.
CC		-I- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC		FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC		PRECEDES ENDOCYTOSIS.
CC		-I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC		specific adenosine on the 28S rRNA.
CC		-I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC		-I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC		CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC		-I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC		INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC		-I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC		or send an email to license@isb-slb.ch).
CC		-----
CC		EMBL; M98344; AAA32624.1; ALT_INT.
DR		EMBL; X54872; -; NOT_ANNOTATED_CDS.
DR		PIR; S32429; TZLSA.
DR		PIR; S24133; S24133.
DR		PDB; 1ABR; 07-FEB-95.
DR		InterPro; IPR001574; RIP.
DR		InterPro; IPR007772; Ricin_B_LectIn.
DR		Pfam; PF00161; RIP; 1.
DR		Pfam; PF00652; Ricin_B_LectIn; 6.
DR		PRINTS; PR00396; SHIGARICIN.
DR		SMART; SMO0458; RICIN; 2.
DR		PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR		PROSITE; PS00275; SHIGA_RICIN; 1.
DR		Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW		Glycoprotein; Lectin; 3D-structure.
KM		
FT	CHAIN	1 251 ABRIN-A CHAIN.
FT	PEPTIDE	252 261 LINKER PEPTIDE.
FT	CHAIN	262 528 ABRIN-B CHAIN.
FT	DOMAIN	273 400 RICIN B-TYPE LECTIN 1.
FT	REPEAT	403 527 RICIN B-TYPE LECTIN 2.
FT	REPEAT	283 325 1-ALPHA.
FT	REPEAT	326 366 1-BETA.
FT	REPEAT	369 401 1-GAMMA.
FT	REPEAT	414 449 2-ALPHA.
FT	REPEAT	453 492 2-BETA.
FT	REPEAT	495 528 2-GAMMA.
FT	ACT_SITE	164 164 BY SIMILARITY.

```

FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).
FT DISULFID 286 305 BY SIMILARITY.
FT DISULFID 329 346 BY SIMILARITY.
FT DISULFID 417 430 BY SIMILARITY.
FT DISULFID 456 473 BY SIMILARITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 298 298 M -> Y (IN REF. 4).
FT CONFLICT 427 427 M -> L (IN REF. 4).
FT CONFLICT 467 467 T -> P (IN REF. 4).
FT CONFLICT 483 483 V -> L (IN REF. 4).
SQ SEQUENCE 528 AA; 59243 MW; ALF76BEC05B9A827 CRC64;

Query Match
Best Local Similarity 24.8%; Score 148; DB 1; Length 528;
Matches 43; Conservative 23; Mismatches 37; Indels 26; Gaps 5;

OY 1 ARFNPIXRRLRQINSSESPPMYMLETSMGRSTQVOQSKDGFNTQIRLQISAGN 60
DB 166 ARFRYSNRVRSIOTGFQFDAMISLENNWMLSKRGVSQDTPPNQ----- 216
OY 61 FVTXSNVRD-----VSSLAIMFECGSRPFSSLDHPSPLLRSVVDANDVTC 109
DB 217 -VTLTINREPIVDLSHPVAVLALMLFCVN--PPNA--NOSPILRSIYEKSK-ICS 270
OY 110 TXSEPTVRI 118
DB 271 SREPTVRI 279

RESULT 5
ABRC_ABRPR STANDARD; PRT: 562 AA.
ID ABRCABRPR
AC P28590:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Abirin-c precursor [Contains: Abirin-c A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); Abirin-c B chain].
OS Eukaryota; Viridiplantae; (Indian licorice) (Crab's eye).
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Leaf;
RC MEDLINE=91266957; PubMed=2050149;
RA Wood K.A., Lord J.M., Mawrzynczak E.J., Platak M.;
RT "Preproabrin: genomic cloning, characterisation and the expression of
RT the A-chain in Escherichia coli.";
RL Eur. J. Biochem. 198;723-732(1991).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA. THE
CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55667; CA939202.1; .
DR PIR; S16022; S16022.
DR HSSP; P11140; 1ABR.
DR InterPro; IPR001574; RIP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00161; RIP; 1.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense: Hydrolase; protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein;
KW Lectin; Signal.
FT SIGNAL 1 34
FT PEPTIDE 35 285
FT CHAIN 286 295
FT DOMAIN 296 562
FT DOMAIN 307 434
FT DOMAIN 437 561
FT REPEAT 317 359
FT REPEAT 360 400
FT REPEAT 403 435
FT REPEAT 448 483
FT REPEAT 487 526
FT REPEAT 529 562
FT ACT_SITE 198 198
FT DISULFID 281 303
FT DISULFID 320 339
FT DISULFID 363 380
FT DISULFID 451 464
FT DISULFID 490 507
FT MOD_RES 35 35
FT CARBOHYD 234 234
FT CARBOHYD 395 395
FT CARBOHYD 435 435
SQ SEQUENCE 562 AA; 62817 MW; 1FD0AB67D7BA6378 CRC64;

Query Match
Best Local Similarity 23.2%; Score 138; DB 1; Length 562;
Matches 43; Conservative 20; Mismatches 50; Indels 6; Gaps 4;

OY 1 ARFNPIXRRLRQINSSESPPMYMLETSMGRSTQVOQSKDGFNTQIRLQISAG 59
DB 200 ARFRYSNRVRSIOTGFQFDAMISLENNWMLSKRGVSQDTPPNVLTSSINQ 259
OY 60 NFVTXSNVRVIVSSLAIMFECGSRPFSSLDHPSPLLRSVVDANDVCTYSEPTVRI 118
DB 260 PVAVDLSHPVAVLALMLFCVN--PPNA--NOSPILRSIYEKSK-ICSREPTVRI 313

RESULT 6
ABRB_ABRPR STANDARD; PRT: 527 AA.
ID ABRB_ABRPR
AC O06077; P81374;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Abirin-b precursor [Contains: Abirin-b A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); Abirin-b B chain].
OS Eukaryota; Viridiplantae; (Indian licorice) (Crab's eye).
OC Spermatophyta; Magnoliophyta; Streptophyta; Tracheophyta;
OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93132798; PubMed=8421313;
RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;

```

RT "Primary structure of three distinct isoabryins determined by cDNA sequencing. Conservation and significance.";
 RT J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 260-527.
 RC TISSUE-Seed;
 RX MEDLINE-93169023; PubMed-7763422;
 RA Kimura M., Sumitazawa T., Funatsu G.;
 RT "The complete amino acid sequences of the B-chains of abrin-a and abrin-b, toxic proteins from the seeds of Abrus precatorius.";
 RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M98345; AAA32625.1; -
 DR HSSP; P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR Pfam; PF00161; RIP; 1.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS02031; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KM plant defense; Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat;
 KM Glycoprotein; Lectin.
 FT CHAIN 1 250
 FT PEPTIDE 251 260
 FT CHAIN 261 527
 FT DOMAIN 272 399
 FT DOMAIN 402 526
 FT REPEAT 282 324
 FT REPEAT 325 365
 FT REPEAT 368 400
 FT REPEAT 413 448
 FT REPEAT 452 491
 FT REPEAT 494 527
 FT ACT_SITE 163 163
 FT DISULFID 246 268
 FT DISULFID 285 304
 FT DISULFID 328 345
 FT DISULFID 416 429
 FT DISULFID 455 472
 FT MOD_RES 1 1
 FT CARBOHYD 110 110
 FT CARBOHYD 360 360
 FT CARBOHYD 400 400
 FT CARBOHYD 282 282
 FT CONFLICT 291 291
 FT CONFLICT 350 351
 FT CONFLICT 378 378
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-> D (IN REF. 2).
 D-> N (IN REF. 2).
 AE-> PO (IN REF. 2).
 SE-> N (IN REF. 2).

FT CONFLICT 426 426 L-> M (IN REF. 2).
 FT CONFLICT 428 428 Y-> D (IN REF. 2).
 FT CONFLICT 431 431 N-> S (IN REF. 2).
 FT CONFLICT 484 484 R-> K (IN REF. 2).
 FT CONFLICT 491 491 N-> S (IN REF. 2).
 FT CONFLICT 493 493 H-> Y (IN REF. 2).
 FT CONFLICT 502 502 R-> G (IN REF. 2).
 FT CONFLICT 509 509 E-> Q (IN REF. 2).
 FT CONFLICT 513 513 H-> W (IN REF. 2).
 FT CONFLICT 516 516 H-> T (IN REF. 2).
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;
 Query Match 22.7%; Score 135; DB 1; Length 527;
 Best Local Similarity 33.3%; Pred. No. 8,7e-08;
 Matches 40; Conservative 24; Mismatches 46; Indels 8; Gaps 5;
 QY 1 AREPNYXWRLRQINSGESSPPNMYWLEETSGRSTQVQSGKDGIFNTQIRLOISGN 60
 DB 165 ARYRFSYRQVGSIRFTNFAFDPAAMISLENWMDNLSGVQOSVODTFPNAVTLR-SYNN 223
 QY 61 --FVTSNVRDVYSSLAIMLFECSGRPFSSLDHPSPLLRVVDANDVTCXSRPTVRI 118
 DB 224 QPVIYDSLTHQSVAVLALMLFCN--PPNA--NQSPLIRSTVERSK-ICSSRYEPTVRI 278
 RESULT 7
 NIGB_SAMNT
 ID NIGB_SAMNT STANDARD; PRT; 563 AA.
 AC P33183; P33184; P93542;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nigirin b precursor (Agglutinin V) (SNV) [contains: Nigirin b A chain (rRNA N-glycosidase) (EC 3.2.2.22); Nigirin b B chain].
 OS Sambucus nigra (European elder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Dipsacales; Adoxaceae; Sambucus.
 OX NCBI_Taxid=4202;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Bark;
 RX MEDLINE-96215449; PubMed-8647092;
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V (Nigirin b), a GalNAC-specific type-2 ribosome-inactivating protein from the bark of elderberry (Sambucus nigra).";
 RT Eur. J. Biochem. 237:505-513(1996).
 RL
 RN [2]
 RP SEQUENCE OF 26-49 AND 298-321.
 RC TISSUE-Bark;
 RX MEDLINE-94003077; PubMed-8400135;
 RA Gilbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R., Munoz R., Atlas F.J., Calonge M., Garcia J.R., Mendez E.;
 RT "Isolation and partial characterization of nigirin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of Sambucus nigra L.";
 RT plant Mol. Biol. 22:1181-1186(1993).
 CC -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF NIGIRIN B TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----


```

DR PRINTS: P000396; SHIGARICIN.
DR PROSITE: P500275; SHIGA_RICIN: 1.
KW Glycoprotein; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 269
FT PROPEP 270 286
FT ACT_SITE 183 183
FT CARBOHYD 250 250
SQ SEQUENCE 266 AA; 31532 MW; E1B013ABEB216CF CRC64;
Query Match 15.4%; Score 92; DB 1; Length 266;
Best Local Similarity 34.1%; Pred. No. 0.0036;
Matches 28; Conservative 16; Mismatches 32; Indels 6; Gaps 3;
QY 1 ARRPPIKMLRQINSGSSPPNMYLLETSMGROSTOYO--QSRDGFINTQIRLOISA 58
DB 185 ARRYIEQOIERARDE--VPSLATISLNSMSGLSKQIOLAGNNGIFRPIVLDNK 242
QY 59 GNEYYSNV--RDVYSLAIME 78
DB 243 GRRVQITNVTSKVTSTNIDLL 264

RESULT 9
RIPA_PHYAM STANDARD; PRT; 294 AA.
AC 003464;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antiviral protein alpha precursor (PAP-1pna) (Ribosome-inactivating
OS Protein) (RNA N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolacaceae; Phytolacca.
OX NCBI_TaxID=3527;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Seed, Leaf, and Root;
RX MEDLINE=93099240; PubMed=1281438;
RA Kataoka J., Habuka N., Masuta C., Miyano M., Koiwai A.;
RT "Isolation and analysis of a genomic clone encoding a pokeweed
RL antiviral protein."
RN Plant Mol. Biol. 20:879-886(1992).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=95010127; PubMed=7925458;
RA Ago H., Kataoka J., Tsuge H., Hakuba N., Inagaki E., Noma M.,
RA Miyano M.;
RT "X-ray structure of a pokeweed antiviral protein, coded by a new
RT genomic clone, at 0.23-nm resolution. A model structure provides a
RT suitable electrostatic field for substrate binding."
RL Eur. J. Biochem. 225:369-374(1994).
CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
CC REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A
CC MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.
CC -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/

```

```

CC or send an email to license@isb.sib.ch).
CC EMBL: D10600; BAA01451.1;
DR PIR: S28421; S28421.
DR PDB: IAPA; 3I-JAN-94.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PROSITE: P500275; SHIGA_RICIN: 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; Cell wall; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 285
FT PROPEP 286 294
FT ACT_SITE 199 199
FT DISULFID 58 282
FT DISULFID 108 130
FT STRAND 28 31
FT HELIX 32 34
FT HELIX 37 51
FT HELIX 57 58
FT STRAND 59 60
FT STRAND 61 63
FT STRAND 67 68
FT STRAND 73 79
FT TURN 81 82
FT STRAND 85 91
FT TURN 92 94
FT TURN 97 104
FT STRAND 105 106
FT TURN 107 113
FT STRAND 114 115
FT TURN 119 129
FT HELIX 126 138
FT STRAND 141 141
FT STRAND 147 154
FT TURN 155 155
FT HELIX 158 160
FT STRAND 163 163
FT HELIX 165 175
FT TURN 176 177
FT TURN 183 195
FT TURN 196 197
FT HELIX 198 202
FT HELIX 204 212
FT TURN 213 215
FT STRAND 218 218
FT STRAND 222 240
FT STRAND 242 242
FT TURN 243 244
FT STRAND 245 253
FT TURN 255 256
FT STRAND 259 264
FT HELIX 265 271
FT STRAND 275 275
SQ SEQUENCE 294 AA; 33069 MW; F2EC27724FA85596 CRC64;
Query Match 14.1%; Score 84; DB 1; Length 294;
Best Local Similarity 26.6%; Pred. No. 0.031;
Matches 21; Conservative 19; Mismatches 37; Indels 2; Gaps 1;
QY 1 ARRPPIKMLRQINSGSSPPNMYLLETSMGROSTOYOQSGDGFINTQIRLOISA 60
DB 201 ARRYIEQOIERARDE--VPSLATISLNSMSGLSKQIOLAGNNGIFRPIVLDNK 258
QY 61 FVYKXSVROVYSSLAIME 79
DB 259 KWIYLRVDIEPDVGLAKT 277

RESULT 10
RIP2_BRVDI STANDARD; PRT; 282 AA.
AC P98184; Q9S8J0;

```



```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin II precursor (rRNA N-
OS glycosidase) (EC 3.2.2.22) (BD2).
OC Bryonia dioica (Red bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Bryonia.
OC NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A.
RA Siegel C.B., Gawlak S.L., Marguardt H.;
RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant
RL Bryonia dioica."
RN Patent number US5597569, 28-JAN-1997.
RP SEQUENCE OF 22-42.
RC TISSUE=ROOT;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RT Marguardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunoconjugates."
RL Biocon. Chem. 5:423-429(1994).
CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; I34238; ; NOT_ANNOTATED_CDS.
DR HSSP; P09989; IMRJ.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.
FT ACT_SITE 183 183 BY SIMILARITY.
FT CARBOHYD 23 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

Query Match 14.0%; Score 83.5; DB 1; Length 282;
Best Local Similarity 31.3%; Pred. No. 0.034; Indels 7; Gaps 4;
Matches 26; Conservative 15; Mismatches 35;

QY 1 ARFNPIXRRLROINSSESPNNMYMLETSMGROSTOVQ-QSKDGIFNTQIRLOI-- 56
DB 185 ARFYIEORVSE--NWGTKRPPDAPFLSLQNMAGSLSEQLQIAQTGEGFARPELVRTVS 242
QY 57 SAGNFVTSXSVRDVYSSIAIMLP 79
DB 243 NTFPEVTNVN-SPVVKIGIALLLY 264

RESULT 11
RIPS_PHYAM STANDARD: PRT; 261 AA.
AC P23339;
DT 01-NOV-1991 (Rel. 20, Created)

```

```

DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antiviral protein S (PAP-S) (Ribosome-inactivating protein) (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.
OC NCBI_TaxID=3527;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=91242096; PubMed=1368643;
RA Kung S.S., Kimura M., Funatsu G.;
RT "The complete amino acid sequence of antiviral protein from the seeds
RT of pokeweed (Phytolacca americana)."
RL Agric. Biol. Chem. 54:3301-3318(1990).
CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
CC SYNTHESIS IN VITRO.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
DR PIR; JE0401; JE0401.
DR HSSP; O03464; IAPA.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin
FT ACT_SITE 175 175 BY SIMILARITY.
FT DISULFD 34 258
FT DISULFD 84 105
SQ SEQUENCE 261 AA; 29200 MW; D88B9962FE8399D CRC64;

Query Match 13.3%; Score 79; DB 1; Length 261;
Best Local Similarity 33.9%; Pred. No. 0.1;
Matches 20; Conservative 12; Mismatches 25; Indels 2; Gaps 1;

QY 1 ARFNPIXRRLROINSSESPNNMYMLETSMGROSTOVQOQSKDGIFNTQIRLOISAG 59
DB 177 ARFYIEQVATKFNPNRDS--PNDKYVLDLENMGKISTAIHNSKNGLPKPELEKNAGD 233

RESULT 12
RIPS_TRIKI STANDARD: PRT; 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein karasurin precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OC NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root tuber;
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT kirilowii var. japonica."
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2]
RP SEQUENCE OF 24-270.
RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,

```

```

RT Karasurin."
RL Chem. Pharm. Bull. 39:1244-1249(1991).
CC -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
CC 60S RIBOSOMAL SUBUNITS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB000666; BAA21786.1;
CC PIR: J00393; J00393.
CC HSSP: P09989; 1MRJ.
CC InterPro: IPR001574; RIP.
CC Pfam: PF00161; RIP. 1.
CC PRINTS: PR00396; SHIGARICIN.
CC PROSITE: PS00275; SHIGA_RICIN: 1.
CC Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
CC KW Toxin; Signal.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT CHAIN 22 270 KARASURIN-C.
CC FT CHAIN 24 270 KARASURIN-A.
CC FT PROPEP 271 289 REMOVED IN MATURE FORM.
CC ACT_SITE 183 183 BY SIMILARITY.
CC FT SEQUENCE 289 AA; 31704 MW; 883D3E242887B26 CRC64;
SQ
Query Match 12.8%; Score 76.5; DB 1; Length 289;
Best Local Similarity 25.3%; Pred. No. 0.22;
Matches 25; Conservative 28; Mismatches 37; Indels 9; Gaps 4;
OY 1 ARENPIYRLRLRQINSGSPPMYMLEETSGRSTGYO--QSDGIFNQLRLQISA 58
DB 185 ARYKFIHQIQIGKRDV--KTLPSLAIISLNSWSALSQIQIASTNNGFETPPVVLINQ 242
OY 59 GNEFTXSNVRD--VISSLAIMLFECSCGRPFSSLDHPSPL 95
DB 243 NORVITINVDAGVYTSNIALIL---NRNNMAALDDVPM 278

```

```

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D21138; BAA04674.1;
CC EMBL: AB022214; BAB09933.1; ALT_SEQ.
CC HSSP: P17119; 3KAR.
CC InterPro: IPR001752; kinesin_motor.
CC Pfam: PF00225; kinesin_1.
CC PRINTS: PR00380; KINESINHEAVY.
CC SMART: SM00129; KISC: 1.
CC PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil;
CC KW Multigene family.
CC FT DOMAIN 1 45 GLOBULAR.
CC FT DOMAIN 46 393 COILED COIL.
CC FT DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
CC FT NP_BIND 481 488 ATP (POTENTIAL).
CC FT SEQUENCE 754 AA; 85030 MW; 76091CD5B5D9C531 CRC64;
SQ
Query Match 12.8%; Score 76; DB 1; Length 754;
Best Local Similarity 50.0%; Pred. No. 0.87;
Matches 15; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
OY 28 ELETSGRSTGYOVSQSDGIFNQLRLQIS 57
DB 315 ELETSSQSTQIRQLDRVNSERLQVS 344

```

```

RESULT 13
KATC_ARATH STANDARD; PRT; 754 AA.
AC P46875; O9FH38;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kinesin-like protein C.
GN KATC OR AT5G54670 OR K5F14.1 OR MRB17.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
OY SEQUENCE FROM N.A.
DB STRAIN=cv Columbia;
RX MEDLINE=94355659; PubMed=8075402;
RA Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
RA Nishikawa K., Takahashi H.;
RT "Sequencing and characterization of the kinesin-related genes katb
RT and katc of Arabidopsis thaliana."
RL Plant Mol. Biol. 25:865-876(1994).

```

```

RESULT 14
RIPA_LUCFY STANDARD; PRT; 277 AA.
AC O00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Luffa.
OX NCBI_TaxID=3670;
OY SEQUENCE FROM N.A.
DB TISSUE=Seed;
RX MEDLINE=92286316; PubMed=1600156;
RA Kataoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from Luffa cylindrica."
RL Plant Mol. Biol. 18:1199-1202(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one

```

```

CC      specific adenosine on the 28S rRNA.
CC      -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC      TYPE 1 RIP SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X62371; CAA44229.1; .
DR      PIR: S22494; S22494.
DR      HSSP: P16094; 1AHC.
DR      InterPro: IPR001574; RIP.
DR      Pfam: PF00161; RIP. 1.
DR      PRINTS: PR00396; SHIGARICIN.
DR      PROSITE: PS00275; SHIGA_RICIN; 1.
KW      Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
FT      Toxin; Signal.
FT      SIGNAL 1 19
FT      CHAIN 20 277
FT      ACT_SITE 179 179
FT      SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;
SQ
Query Match
Best Local Similarity 12.4%; Score 74; DB 1; Length 277;
Matches 21; Conservative 18; Mismatches 37; Indels 4; Gaps 2;
OY      1 AARPNTXWMLRQINSSESPNNMYMLETSWGRSTQVQ--OSKDGIPNTQIRLOISA 58
DB      181 SRFKYIEGIIERIS--KNQVPSLATISLENWMSLSKOILAQTNNGTFTPVVITDDK 238
OY      59 GNEVTSXNVNRYVSSLAIML 78
DB      239 GQRETVNTVTSKVTKNIQL 258
RESULT 15
RIP1_CUCFI STANDARD; PRT; 286 AA.
AC      Q9FRX4;
DR      16-OCT-2001 (Rel. 40, Created)
DR      16-OCT-2001 (Rel. 40, Last sequence update)
DR      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)
DE      (EC 3.2.2.22).
OS      Cucumis figareli.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX      NCBI_TaxID=131071;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Yamada T., Ohki S.T., Osaka T.;
RT      "Cloning and analysis of a cDNA coding a putative ribosome-
RT      inactivating protein from Cucumis figareli.";
RL      Plant Biotechnol. 17:337-340(2000).
CC      -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC      specific adenosine on the 28S rRNA.
CC      -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC      TYPE 1 RIP SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).

```

```

DR      EMBL: AB045560; BAB19677.1; .
DR      HSSP: P16094; 1AHC.
DR      InterPro: IPR001574; RIP.
DR      Pfam: PF00161; RIP. 1.
DR      PRINTS: PR00396; SHIGARICIN.
DR      PROSITE: PS00275; SHIGA_RICIN; 1.
KW      Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
FT      Toxin; Signal.
FT      SIGNAL 1 21
FT      CHAIN 22 286
FT      ACT_SITE 185 185
FT      CARBOHYD 103 103
FT      CARBOHYD 110 110
FT      CARBOHYD 252 252
SQ      SEQUENCE 286 AA; 31771 MW; 4EFD4966E04DAA1 CRC64;
Query Match
Best Local Similarity 12.4%; Score 74; DB 1; Length 286;
Matches 18; Conservative 13; Mismatches 29; Indels 2; Gaps 1;
OY      19 SSPPNNMYMLETSWGRSTQVQ--OSKDGIPNTQIRLOISA 76
DB      203 SKVPDLAALSLENWMSLSKOIAKSNNGQPTQVVKIINDKGLTEVTNVSLSVTKNI 262
OY      77 ML 78
DB      263 ML 264

```

Search completed: March 18, 2003, 08:25:43
Job time : 12.3819 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:18 ; Search time 15.8467 Seconds

(without alignments)
1547.297 Million cell updates/sec

Title: US-09-627-165E-14

Perfect score: 596
Sequence: 1 ARNPPIXRRLRQINGSSRSS.....VYDANDVYCTXSEPTVRIV 119

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	408.5	68.5	565	10 Q8W243	Q8W243 viscum albu
2	387	64.9	531	10 Q8RXH6	Q8RXH6 viscum albu
3	324	54.4	449	10 Q8RXH7	Q8RXH7 viscum albu
4	193	32.4	592	10 Q8W2E7	Q8W2E7 iris hollan
5	187	31.4	573	10 Q8W2E8	Q8W2E8 iris hollan
6	179	30.0	549	10 Q9FV22	Q9FV22 cinamomum
7	179	30.0	580	10 Q94BW4	Q94BW4 cinamomum
8	179	30.0	580	10 Q94BW3	Q94BW3 cinamomum
9	179	30.0	581	10 Q94BW5	Q94BW5 cinamomum
10	178	29.9	541	10 Q41174	Q41174 ricinus com
11	157	26.3	382	10 Q8SMA3	Q8SMA3 abrus preca
12	139	23.3	528	10 Q06076	Q06076 abrus preca
13	128.5	21.6	603	10 Q9M653	Q9M653 polygonatum
14	128.5	21.6	604	10 Q9M654	Q9M654 polygonatum
15	122	20.5	547	10 Q9M6E9	Q9M6E9 abrus preca
16	121.5	20.4	570	10 Q41358	Q41358 sambucus n1

17	119.5	20.1	570	10 Q22415	Q22415 sambucus n1
18	109	18.3	563	10 Q04367	Q04367 sambucus n1
19	109	18.3	564	10 Q9AVR2	Q9AVR2 sambucus eb
20	107.5	18.0	251	10 Q96236	Q96236 abrus preca
21	107.5	18.0	252	10 Q38761	Q38761 abrus preca
22	107	18.0	563	10 Q94552	Q94552 sambucus n1
23	103	17.3	565	10 Q04071	Q04071 sambucus n1
24	103	17.3	566	10 Q04072	Q04072 sambucus n1
25	100.5	16.9	251	10 Q96235	Q96235 abrus preca
26	100.5	16.9	251	10 Q96237	Q96237 abrus preca
27	100.5	16.9	569	10 P93543	P93543 sambucus n1
28	98	16.4	300	10 Q04356	Q04356 iris hollan
29	95	15.9	298	10 Q04358	Q04358 iris hollan
30	93.5	15.7	252	10 Q38760	Q38760 abrus preca
31	91	15.3	300	10 Q04357	Q04357 iris hollan
32	88	14.8	293	10 Q84552	Q84552 jatropa cu
33	85	14.3	293	10 Q8VYU0	Q8VYU0 jatropa cu
34	81	13.6	313	10 Q941G8	Q941G8 phycolacca
35	81	13.6	339	10 Q8RTA4	Q8RTA4 phycolacca
36	79	13.3	261	10 Q8S947	Q8S947 phycolacca
37	77.5	13.0	289	10 Q41216	Q41216 trichosanthe
38	75	12.6	279	10 Q9AUE3	Q9AUE3 chenopodium
39	75	12.6	286	2 Q50196	Q50196 mycobacteri
40	75	12.6	302	10 P93077	P93077 clerodendru
41	74.5	12.5	556	4 Q9NV60	Q9NV60 homo sapien
42	74.5	12.5	1023	4 Q5ULF2	Q5ULF2 homo sapien
43	72.5	12.2	289	10 Q94KE4	Q94KE4 trichosanthe
44	72	12.1	144	16 Q93GP4	Q93GP4 salmonele
45	72	12.1	237	10 Q9ATB3	Q9ATB3 phycolacca

ALIGNMENTS

RESULT 1

ID Q8W243 PRELIMINARY; PRT: 565 AA.
AC Q8W243;
DT 01-MAR-2002 (TrEMBLrel. 20, created)
DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
DE VCA precursor.
OS viscum album subsp. coloratum
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=159976;
RN [1]
RP SEQUENCE FROM N.A.
RA Park W.-B., Lyu S.;
RT Cloning of Viscum album subsp. coloratum (Korean mistletoe).
RI Biochem. Biophys. Res. Commun. 0:0-0(2002).
DR EMBL: AF369951; AAL0417.1.
DR InterPro: IPR001898; Gfam_pos_anchor.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 273 VCA ALPHA CHAIN.
FT CHAIN 309 565 VCA BETA CHAIN.
SQ SEQUENCE 565 AA: 62401 MW: 9915394DaO5F11 CRG64;

Query Match 68.5%; Score 408.5; DB 10; Length 565;
Best Local Similarity 69.7%; Pred. No. 1,6e-38;
Matches 85; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

```

||||| 11 : ||||| : ||||| ||||| 11 : || 11 ||
DB 200 ARFNPWFARARQYINSGESFLPDMYMLETETSMGQOSTQVOQSTDVGFNNPFLIGSTGN 259
QY 61 FVTXSNVNDVYSSLAIMLFECGSRPSSLDHPSPLLSVVD---ANDVTCXSEPTVR 117
   ||| ||||| : ||||| : ||| 11 : || : ||||| |||||
DB 260 FVTLSNVNDVYASLAIMLFVCGRDRPSSDVRWYPLVIRPVLENGAAYDVTCIASSEPTVR 319
QY 118 IV 119
   ||
DB 320 IV 321

RESULT 2
Q8RXH6 PRELIMINARY: PRT: 531 AA.
AC Q8RXH6:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lectin chain A isoform 1 (fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramesivam M., Misra V., Srinivasan A., Singh T.P.;
RT "Viscum album (Indian) mRNA for Mistletoe Lectin chain A isoform 1 and
   chain B."
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
EMBL; AY081149; AA187006.1; -.
FT NON_TER 1 249
FT CHAIN 1 249 LECTIN CHAIN A ISOFORM 1.
FT CHAIN 269 531 LECTIN B CHAIN.
SQ SEQUENCE 531 AA: 58802 MW: 18244BEFEF35422 CRC64;

Query Match 64.9%; Score 387; DB 10; Length 531;
Best Local Similarity 67.2%; Pred. No. 4,5e-36;
Matches 80; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

QY 1 ARFNPIKWLRLROINSGESPPNMYMLETETSMGQOSTQVOQSKDGFNTQIRLOISAGN 60
   ||||| 11 : ||||| : ||||| ||||| 11 : || 11 ||
DB 167 ARFNPIKWLRLROINSGESFLPDMYMLETETSMGQOSTQVOQSTREGVFNPIRLAIPGN 226
QY 61 FVTXSNVNDVYSSLAIMLFECGSRPSSLDHPSPLLSVVDANDVTCXSEPTVR 119
   ||| ||||| : ||||| : ||| 11 : || : ||||| |||||
DB 227 FVTLSNVNDVYASLAIMLFVCGRDRPSSDVRWYPLVIRPV--ADVTCIASSEPTVR 283

RESULT 3
Q8RXH7 PRELIMINARY: PRT: 249 AA.
AC Q8RXH7:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lectin chain A isoform 2 (fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramesivam M., Srinivasan A., Singh T.P.;
RT "Viscum album (Indian) mRNA for Mistletoe Lectin Chain A, isoform 2."
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
EMBL; AY081148; AA187005.1; -.
FT NON_TER 1 249
FT CHAIN 1 249

```

```

SQ SEQUENCE 249 AA: 27944 MW: 89FAFB78309A83B3 CRC64;

Query Match 54.4%; Score 324; DB 10; Length 249;
Best Local Similarity 77.8%; Pred. No. 3,4e-29;
Matches 63; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 ARFNPIKWLRLROINSGESPPNMYMLETETSMGQOSTQVOQSKDGFNTQIRLOISAGN 60
   ||||| 11 : ||||| : ||||| ||||| 11 : || 11 ||
DB 167 ARFNPIKWLRLROINSGESFLPDMYMLETETSMGQOSTQVOQSTDVGFNNPFLIGSTGN 226
QY 61 FVTXSNVNDVYSSLAIMLFEC 81
   ||| ||||| : ||||| : ||| 11 : || : ||||| |||||
DB 227 FVTLSNVNDVYASLAIMLFVC 247

RESULT 4
Q8W2E7 PRELIMINARY: PRT: 592 AA.
AC Q8W2E7:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ribosome-inactivating protein IRAB.
GN LECIRAB.
OS Iris hollandica (Dutch iris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OC Iris.
OX NCBI_TaxID=35876;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Damme E.J.M., Peumans W.J.;
RT "Iris (Iris hollandica var. Professor Blaauw) plants express both type
   1 and type 2 ribosome-inactivating proteins in bulb tissue."
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
EMBL; AF256085; AA155094.1; -.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN_2.
DR PROSITE: PS0231; RICIN; LECTIN; 2.
DR PROSITE: PS00275; SHIGA; RICIN; UNKNOWN; 1.
SQ SEQUENCE 592 AA: 65677 MW: 6FLF335DCEFA99E3 CRC64;

Query Match 32.4%; Score 193; DB 10; Length 592;
Best Local Similarity 37.7%; Pred. No. 1,2e-13;
Matches 52; Conservative 20; Mismatches 42; Indels 24; Gaps 6;

QY 1 ARFNPIKWLRLR---QINSGESPPNMYMLETETSMGQOSTQVOQSKDGFNTQIRLOIS 57
   ||| 11 : ||| 11 : ||||| ||||| 11 : || 11 ||
DB 210 ARFAIETFRVRNRNIQVGYRSFRGAGMLDETNNWGTLSERVOESNECVFANRLITDTT 269
QY 58 AGNEFVT---XSNVNDVYSSLAIMLFECGSRPSSLD-----HPSDL-----LRSV 100
   ||| 11 : ||| 11 : ||||| ||||| 11 : || 11 ||
DB 270 --NEFTIHITNAQTARQVCGMALILFACKAR--QSLQALPPHDSVLPITLIDLNVVRS 325
QY 101 VDAANDVTCXSEPTVR 118
   : ||| 11 : ||||| : ||| 11 : ||| 11 : ||
DB 326 LDYVEDTCTVSEPTVR 343

RESULT 5
Q8W2E8 PRELIMINARY: PRT: 573 AA.
AC Q8W2E8:
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ribosome-inactivating protein IRAR (fragment).
GN LECTIRAR.
OS Iris hollandica (Dutch iris).

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OX NCBI_TaxID=35876;
RN
RP SEQUENCE FROM N.A.
RA Van Damme E.J.M., Peumans W.J.;
RT "Iris (Iris hollandica var. Professor Blaauw) plants express both type
RT 1 and type 2 ribosome-inactivating proteins in bulb tissue.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF256084; AAL55093.1; -
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PRO0396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 573 AA; 63759 MW; 1414A39AECDAF5C CRC64;

Query Match 31.4%; Score 187; DB 10; Length 573;
Best Local Similarity 37.7%; Pred. No. 5.6e-13;
Matches 52; Conservative 19; Mismatches 43; Indels 24; Gaps 6;

QY 1 ARNPXWRLRR--QINSGSSPPNMYMLETSMGROSTOVQOQSDGIFNQIRQIS 57
DB 191 ARFALFRNRNIIQVGDYRSPFGAGMIDLETNMEPLSERQESNEGAFANRLTQTS 250
QY 58 AGNRYT---XSNRVYISSLAIMLFECSSGRPFSSLD-----HSPPL-----LRSV 100
DB 251 --NETIDIVNAQYARQVCGALLLFACKAR--QSLQALPRLPHSPVLPITLLNVVRS 306
QY 101 VDAANDVCTXSEPTVRI 118
DB 307 LDIVEDDTCPLEPTMRI 324

RESULT 6
Q9FV22 PRELIMINARY; PRT; 549 AA.
ID 09FV22
AC 09FV22;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE rRNA-glycosidase (EC 3.2.2.22) (Fragment).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN
RP SEQUENCE FROM N.A.
RA Xie L., Liu W.-Y., Wang E.-D.;
RT "Molecular cloning of cinnamomin A", B-chain and the expression,
RT purification, characterization and mutagenesis of the A-chain.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
CC EMBL; AF259548; AAF68978.2; -.
DR HSSP; P02879; 2AA1.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 5.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PRO0396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR Hydroxylase; Toxin.
FT NON_TER 1
SQ SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;

```

```

Query Match 30.0%; Score 179; DB 10; Length 549;
Best Local Similarity 38.9%; Pred. No. 4.5e-12;
Matches 51; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 RNPXWRLRRQINSGSSPPNMYMLETSMGROSTOVQOQSDGIFNQIRQIS 60
DB 170 RNFETRYRVRGSISSRAEMRPDPAMLSLEKMSALSNAVOQSGVSSPVELSISNK 229
QY 61 FVYXSNVRD-VISSLAIMLFECSSGRPFSS-----LDHSPPLLRVY-----DAANDV 107
DB 230 PYYVGSVSDRVISGLAIMLFICRSDRASSDQFIDH--MLMIRPLIVDAVATDANDD 287
QY 108 TCTXSEPTVRI 118
DB 288 TCADPEPTVRI 298

RESULT 7
Q94BW4 PRELIMINARY; PRT; 580 AA.
ID 094BW4
AC 094BW4;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE rRNA-glycosidase precursor (EC 3.2.2.22).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.-Z., Liu W.-Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
CC EMBL; AF039802; AAK82459.1; -.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 580 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN II.
SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECC0CBFF CRC64;

Query Match 30.0%; Score 179; DB 10; Length 580;
Best Local Similarity 40.5%; Pred. No. 4.8e-12;
Matches 53; Conservative 16; Mismatches 46; Indels 16; Gaps 6;

QY 2 RNPXWRLRRQINSGSSPPNMYMLETSMGROSTOVQOQSDGIFNQIRQIS 60
DB 202 RNFETRYRVRGSISSRAEMRPDPAMLSLEKMSALSNAVOQSGVSSPVELSISNK 261
QY 61 FVYXSNVRD-VISSLAIMLFEC--SGRPFSS--LDHSPPLLRVY-----DAANDV 107
DB 262 PYYVGSVSDRVISGLAIMLFICRSDRRTSSDQFIDH--LLMIRPLIVDAVATDANDD 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADPEPTVRI 330

RESULT 8
Q94BW3 PRELIMINARY; PRT; 580 AA.
ID 094BW3
AC 094BW3;
DT 01-DEC-2001 (Tremblrel. 19, Created)

```

```

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE rRNA -glycosidase precursor (EC 3.2.2.22).
OS Cinamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu M.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinamomnin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039803; AAK82460.1; -.
DR InterPro: IPR000772; RIBCIN_B_lectin.
DR Pfam: PF00652; RIBCIN_B_lectin; 6.
DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 580 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7EB558 CRC64;

Query Match 30.0%; Score 179; DB 10; Length 580;
Best Local Similarity 38.9%; Pred. No. 4.8e-12;
Matches 51; Conservative 17; Mismatches 47; Indels 16; Gaps 5;

QY 2 RBNPIXWLRRLROINSSESPPMYMLEETSWGROSTOVQOS-KDGIPTQIRLQISAGN 60
DB 202 RFRLEFYRRESITRAEMFRPDPALSLKMKWALSNAVOOSNGGVFSPELRSISNK 261
QY 61 FVTXSNVRD-VTSSLAIMEFGSGRPFSS---LDHPSLLRLRSV-----DAANDY 107
DB 262 PTVGVSVDRIYSLAIMEFLICRSTDRASSDQFIDH--LIMIRPLDVAVATDADND 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADEPTVRI 330

RESULT 9
Q94BW5 PRELIMINARY; PRT; 581 AA.
AC Q94BW5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE rRNA -glycosidase precursor (EC 3.2.2.22).
OS Cinamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu M.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinamomnin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039801; AAK82458.1; -.
DR InterPro: IPR000772; RIBCIN_B_lectin.
DR Pfam: PF00652; RIBCIN_B_lectin; 5.
DR Pfam: PF00161; RIP; 1.

```

```

DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 581 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN I.
SQ SEQUENCE 581 AA; 64215 MW; 6EBF5FB8FBA3D196 CRC64;

Query Match 30.0%; Score 179; DB 10; Length 581;
Best Local Similarity 38.9%; Pred. No. 4.8e-12;
Matches 51; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 RBNPIXWLRRLROINSSESPPMYMLEETSWGROSTOVQOS-KDGIPTQIRLQISAGN 60
DB 202 RFRLEFYRRESITRAEMFRPDPALSLKMKWALSNAVOOSNGGVFSPELRSISNK 261
QY 61 FVTXSNVRD-VTSSLAIMEFGSGRPFSS---LDHPSLLRLRSV-----DAANDY 107
DB 262 PTVGVSVDRIYSLAIMEFLICRSTDRASSDQFIDH--LIMIRPLDVAVATDADND 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADEPTVRI 330

RESULT 10
Q41174 PRELIMINARY; PRT; 541 AA.
AC Q41174;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
OS Ricinus communis (Castor bean).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OX eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3986;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92338377; Pubmed-1633311;
RA Roberts L.M., Tregear J.W., Lord J.M.;
RT "Molecular cloning of ricin.";
RL Targeted Diagn. Ther. 7:81-97(1992).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: S40366; AAB22582.1; -.
DR HSSP: P02879; IBR6.
DR InterPro: IPR001572; RIBCIN_B_lectin.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00652; RIBCIN_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RIBCIN_2.
DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; UNKNOWN_1.
KW Hydrolase; Toxin.
FT NON TER 1 1
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1FE29D9 CRC64;

Query Match 29.9%; Score 178; DB 10; Length 541;
Best Local Similarity 37.0%; Pred. No. 5.7e-12;
Matches 44; Conservative 19; Mismatches 52; Indels 4; Gaps 1;

QY 1 ARBNPIXWLRRLROINSSESPPMYMLEETSWGROSTOVQOSKDGIPTQIRLQISAGN 60
DB 179 ARFOYLEGEKTRIRIRNRRSAPDPVITLLENSWGRSLTAIOESNGCAFASPIQLORRNS 238
QY 61 FVTXSNVRD-VTSSLAIMEFGSGRPFSSLDHPSLLRLRSVVAANDVCTXSEPTVRI 119
DB 239 KFSYDVSIPLITLALMYKACAPPSSQPS---LIRVVPNFNADVCMDEPTVRI 293

```


OY 105 -----NDYCTXSEPTVRI 118
 DB 325 AGEYDEFTCTVGEPTRRI 344

RESULT 14

O9M654 PRELIMINARY; PRT; 604 AA.
 AC O9M654
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA_glycosidase (EC 3.2.2.22).
 GN RIPM.
 OS Polygonatum multiflorum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
 OC Convallariaceae; Polygonatum.
 OK NCBI_TaxID=45371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20247177; PubMed=10785398;
 RA Van Damme E.J., Hao Q., Charels D., Barre A., Rouge P., Van Leuven F.,
 Peumans W.J.:
 RT "Characterization and molecular cloning of two different type 2
 RT ribosome-inactivating proteins from the monocotyledonous plant
 RT Polygonatum multiflorum."
 RL Eur. J. Biochem. 267:2746-2759(2000).
 CC -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: AF213983; AAF37218.1;
 DR HSSP: P02879; 2AAI.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00652; Ricin_B_lectin; 5.
 DR PRINTS: PR00161; RIP; 1.
 DR PROSITE: PS00396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE: PS00275; SHIGA_RICIN; UNKNOWN_1.
 KM Hydrolase; Toxin.
 SQ SEQUENCE 604 AA; 67060 MW; 6206C2B25F90804E CRC64;

Query Match 21.6%; Score 128.5; DB 10; Length 604;
 Best Local Similarity 30.9%; Pred. No. 3.3e-06;
 Matches 42; Conservative 21; Mismatches 48; Indels 25; Gaps 6;

OY 1 ARFNPIMRLRQINSSESPNNMYLETSWGRSTOVQOSKD-----GIFNTQIRL 54
 DB 213 ARFRRIEVOYRSIADDDTFPGSLMTMEKKMSQOVERSVNDGIFGTFTVQL 272
 OY 55 -----QISAGNEVYXSNVDYISLAIMLFECGSRFSSLDHSP-----LLKRSVD 102
 DB 273 IDNMLQTLINDNF-----NALSLHTMLAILFRC--RTTSSHMTLPASNIIVLMGEDYD 326
 OY 103 AANDVCTXSEPTVRI 118
 DB 327 -KDDEKCTVGEPTRRI 341

RESULT 15

O9M6E9 PRELIMINARY; PRT; 547 AA.
 AC O9M6E9;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA_glycosidase (EC 3.2.2.22).
 GN AAG.
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euphorbia; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 OK NCBI_TaxID=3816;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=20102702; PubMed=10636890;
 RA Liu C.L., Tsai C.C., Lin S.C., Wang L.I., Hsu C.I., Hwang M.J.,
 Lin J.Y.:
 RT "Primary Structure and Function Analysis of the Abrus precatorius
 RT Agglutinin A Chain by Site-directed Mutagenesis: Prolong of Amphiphilic
 RT alpha-Helix H Impacts Protein Synthesis Inhibitory Activity";
 RL J. Biol. Chem. 275:1897-1901(2000).
 CC -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: AF190173; AAF28309.1;
 DR HSSP: P1140; 1ABR.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00652; Ricin_B_lectin; 6.
 DR PRINTS: PR00161; RIP; 1.
 DR PROSITE: PS00396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
 DR PROSITE: PS00275; SHIGA_RICIN; UNKNOWN_1.
 KM Hydrolase; Toxin.
 SQ SEQUENCE 547 AA; 61248 MW; 355A325C2354A1BD CRC64;

Query Match 20.5%; Score 122; DB 10; Length 547;
 Best Local Similarity 31.9%; Pred. No. 1.7e-05;
 Matches 38; Conservative 22; Mismatches 53; Indels 6; Gaps 4;

OY 1 ARFNPIMRLRQINSSESPNNMYLETSWGRSTOVQOSKDCIFNTQIRL-QISAG 59
 DB 185 ARFRVYSKLVISLSNRAAFQDPDSMLSEMTWEPISRVAOHTVDTFPQNTLINVROE 244
 OY 60 NEVYXSNVDYISLAIMLFECGSRFSSLDHSPPLLSRVYDANDVCTXSEPTVRI 118
 DB 245 RRVYSSLSHPSVALMLFVCN--PLNATQ--SPILLINSVVEQSK-ICSSHREPTVRI 298

Search completed: March 18, 2003, 08:26:48
 Job time : 20.8467 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:00:53 ; Search time 49.0704 Seconds

(Without alignments)
757.625 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425

Sequence: 1 HQTGDEYFRITLRLDYVS.....NSGAVDVTCTASEPTVRLIV 279

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1389	97.5	551	23	ABR79450
2	1269.5	89.1	531	20	AAV25979
3	1269.5	89.1	532	20	AAV25982
4	1269.5	89.1	564	18	AAW10021
5	1269.5	89.1	564	20	AAW90127
6	1211	85.0	533	20	AAV25970
7	1211	85.0	533	20	AAV25973
8	1211	85.0	533	20	AAV25976
9	1155	81.1	254	22	AAW47090
10	1152	80.8	256	20	AAV25981

11	1152	80.8	256	20	AAV25984
12	1131	79.4	252	19	AAW64661
13	1131	79.4	253	18	AAW10022
14	1131	79.4	253	20	AAW90125
15	1127	79.1	252	19	AAW64659
16	1119	78.5	254	20	AAV25980
17	1119	78.5	254	20	AAV25983
18	1090	76.5	254	22	AAW47091
19	1080.5	75.8	255	20	AAV25971
20	1080.5	75.8	255	20	AAV25974
21	1080.5	75.8	255	20	AAV25977
22	1067	74.9	256	22	AAW47092
23	885	62.1	174	22	AAW47097
24	474	33.3	251	21	AAW19267
25	466.5	32.7	250	18	AAW25137
26	466.5	32.7	250	18	AAW21700
27	461	32.4	252	16	AAW81976
28	458	32.1	252	16	AAW81977
29	452	31.7	253	16	AAW81978
30	425.5	29.9	332	8	AAW70097
31	425.5	29.9	332	8	AAW70098
32	425.5	29.9	332	10	AAW95839
33	425.5	29.9	332	11	AAW06554
34	425.5	29.9	562	10	AAW90079
35	425.5	29.9	576	8	AAW70326
36	425.5	29.9	576	18	AAW25787
37	425.5	29.9	576	20	AAV55892
38	425.5	29.9	576	21	AAV78592
39	425.5	29.9	576	22	AAW78301
40	425.5	29.9	576	22	AAW78302
41	423.5	29.7	574	10	AAW94793
42	423	29.7	540	18	AAW25143
43	423	29.7	540	18	AAW21706
44	422.5	29.6	565	6	AAW50166
45	422.5	29.6	565	7	AAW60240

ALIGNMENTS

RESULT 1	
ABR79450	
ID	ABR79450 standard; Protein; 551 AA.
XX	
AC	ABR79450:
XX	
DT	08-JUL-2002 (first entry)
XX	
DE	Galactose-recognising mistletoe lectin.
XX	
KW	Mistletoe; galactose-recognising mistletoe lectin; MLIII.
XX	
OS	Viscum album.
XX	
FT	Key
FT	Misc-difference 223
FT	/note= "Encoded by ATG"
FT	Misc-difference 251
FT	/note= "Encoded by TTP"
FT	Misc-difference 344
FT	/note= "Encoded by TCC"
FT	Misc-difference 380
FT	/note= "Encoded by GCC"
FT	Misc-difference 448
FT	/note= "Encoded by GTG"
FT	
PN	DE10044027-A1.
XX	
PD	14-MAR-2002.
XX	
PF	06-SEP-2000; 2000DE-1044027.
XX	
PR	06-SEP-2000; 2000DE-1044027.

Mistletoe lectin A
Mistletoe rMLA var
Prepro mistletoe I
Mistletoe ML A-cha
Mistletoe rMLA pro
Mistletoe lectin A
Mistletoe lectin A
A-chain isoform fo
Mistletoe lectin A
Mistletoe lectin A
Mistletoe lectin A
A-chain isoform fo
Amino acid sequenc
Abirin A A-chain
Abirin A A-chain
Recombinant abirin
Recombinant abirin
Riclin A. Escheric
Sequence of Riclinu
Riclin A encoded by
Riclin A gene produ
Riclin D. Riclinu
Sequence of Riclinu
Castorbean riclin.
Castor bean riclin
Riclinus communis r
Castor bean prepro
DNA sequence of ri
Castor oil plant A
R. communis agglut
Sequence of prepro
PreproRiclin. Ricl

```
XX PA (VISC-) VISCUM AG.
XX PI Kieff S;
XX DR WPI: 2002-316737/36.
XX DR N-PSDB; ABL56947.
XX PT New nucleic acid encoding preprotein of mistletoe lectin, useful as
XX PT diagnostic and therapeutic agents, also encodes polypeptide
XX PS Claim 1; Fig 1; 6pp; German.
XX CC The invention relates to a nucleic acid molecule (ABL56947) that encodes
XX CC a preprotein (AB879450) which, after maturation, has the biological
XX CC activity of the galactose-recognising mistletoe lectin (MLII). The MLII
XX CC encoding nucleic acid molecule, primers specific to it or complements of
XX CC it, and encoded (oligomeric) polypeptides are useful as diagnostic and
XX CC therapeutic agents.
SQ Sequence 551 AA;
Query Match 97.5%; Score 1389; DB 23; Length 551;
Best Local Similarity 97.5%; Pred. No. 1.6e-133;
Matches 272; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLRSTIPVSDAQRVYVELTNGGDSITPA 60
DB 43 HQTGGEYFRFTLLRDYVSSGFSNEIPLRSTIPVSDAQRVYVELTNGGDSITPA 102
OY 61 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTRSSLPFTGTYDLERYAGHRDQIP 120
DB 103 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTRSSLPFTGTYDLERYAGHRDQIP 162
OY 121 LGIEELIOSVSAALRYGSGSTRQAARSLITLQMISEAARFNPFMRARQYINSGESFLPD 180
DB 163 LGIEELIOSVSAALRYGSGSTRQAARSLITLQMISEAARFNPFMRARQYINSGESFLPD 222
OY 181 MYMLELETSWGOOSTOVQOSTDGVFNPNPRLGISTGNFVTLSNVRDVIASLAIMLFVCRD 240
DB 223 VYMLELETSWGOOSTOVQOSTDGVFNPNPRLAISGNEFTLTSNVRDVIASLAIMLFVCRD 282
OY 241 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 279
DB 283 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 321
RESULT 2
AAV25979
ID AAV25979 standard; Protein; 531 AA.
XX AC AAV25979;
XX DT 18-OCT-1999 (first entry)
XX DE Mistletoe lectin I protein fragment.
XX KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
XX KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KM cancer; cytotoxicity; antigen; isoform; lectin I.
XX OS Viscum album.
XX PN DEL19804210-A1.
XX PD 12-AUG-1999.
XX PF 03-FEB-1998; 98DE-1004210.
XX PR 03-FEB-1998; 98DE-1004210.
XX PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
```

```
XX PI Morris P, Stiefel T, Voelter W, Welters P;
XX DR WPI: 1999-445335/38.
XX DR N-PSDB; AAZ09103.
XX PT Preparation of mistletoe lectins in heterologous systems,
XX PT particularly for use as anticancer agents and immunostimulants
XX PS Claim 7; Fig 1B; 78pp; German.
XX CC This invention describes a novel mistletoe lectin (I) and its fragments
XX CC which have antitumour and immunostimulatory activity. The A-chain (MIA)
XX CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX CC fragments are used to treat uncontrolled cell growth (particularly
XX CC cancers) and if they lack cytotoxicity, to increase the strength of the
XX CC immune response, particularly to a co-administered antigen
XX CC (tumour-associated, bacterial or viral). The method allows production of
XX CC mistletoe lectin, and its individual chains, in many different isoforms
XX CC and on a large scale, at any time of the year. Recombinant products are
XX CC free from toxins present in natural mistletoe extracts. This sequence
XX CC represents a mistletoe lectin I protein fragment.
SQ Sequence 531 AA;
Query Match 89.1%; Score 1269.5; DB 20; Length 531;
Best Local Similarity 88.9%; Pred. No. 2.8e-121;
Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
OY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLRSTIPVSDAQRVYVELTNGGDSITPA 60
DB 10 HQTGGEYFRFTLLRDYVSSGFSNEIPLRSTIPVSDAQRVYVELTNGGDSITPA 69
OY 61 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTRSSLPFTGTYDLERYAGHRDQIP 120
DB 70 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTRSSLPFTGTYDLERYAGHRDQIP 129
OY 121 LGIEELIOSVSAALRYGSGSTRQAARSLITLQMISEAARFNPFMRARQYINSGESFLPD 180
DB 130 LGIDQLIOSVYALRFPGSGSTRQAARSLITLQMISEAARFNPFMRARQYINSGESFLPD 189
OY 181 MYMLELETSWGOOSTOVQOSTDGVFNPNPRLGISTGNFVTLSNVRDVIASLAIMLFVCRD 240
DB 190 VYMLELETSWGOOSTOVQOSTDGVFNPNPRLAIPGAFVTLTNVRDVIASLAIMLFVCGE 249
OY 241 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 279
DB 250 RPSSSDVRYWPLVIRPVI-----ADVTGCSASEPTVRIY 283
RESULT 3
AAV25982
ID AAV25982 standard; Protein; 532 AA.
XX AC AAV25982;
XX DT 18-OCT-1999 (first entry)
XX DE Mistletoe lectin I (variant) protein fragment.
XX KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
XX KM ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KM cancer; cytotoxicity; antigen; isoform; lectin I.
XX OS Viscum album.
XX PN DEL19804210-A1.
XX PD 12-AUG-1999.
```

PF 03-FEB-1998; 98DE-1004210.
 XX 03-FEB-1998; 98DE-1004210.
 PR (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 PA
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI, 1999-445335/38.
 DR N-PSDB; AAO09106.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure; Fig 4B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin I protein variant.
 XX
 SQ Sequence 532 AA;
 Query Match 89.1%; Score 1269.5; DB 20; Length 532;
 Best Local Similarity 88.9%; Pred. No. 2.8e-121;
 Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
 QY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAORFVLELTNOGDSITAA 60
 Db 10 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAORFVLELTNOGDSITAA 69
 QY 61 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLEFTGTRSSLPFTGSYTLDERYAGHRDQIP 120
 Db 70 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLEFTGTRSSLPFTGSYTLDERYAGHRDQIP 129
 QY 121 LGIEELIQSVSALRYPGGSTRQAARSLILILQIMISEAARFNPIFWARQYINSGESFLPD 180
 Db 130 LGIDQLIOSYVTLRFPGGSTRQAARSLILILQIMISEAARFNPIFWARQYINSGESFLPD 189
 QY 181 MYMLEETSMGQOSTVOVQOSTDGVFNPFRLGISTGNFVTLISNVBDVYIASLAIMLFCRD 240
 Db 190 VYMLEETSMGQOSTVOVQOSTDGVFNPFRLGISTGNFVTLISNVBDVYIASLAIMLFCRD 249
 QY 241 RPSSSDVRYWPLYIRPVLENLSCAVDDYCTASEPTVRIY 279
 Db 250 RPSSSDVRYWPLYIRPVYI-----ADVTCSASEPTVRIY 283
 RESULT 4
 AAM10021
 ID AAM10021 standard; Protein; 564 AA.
 XX
 AC AAM10021;
 XX
 DT 18-DEC-1997 (first entry)
 XX
 DE Prepro mistletoe lectin.
 XX
 KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.
 XX
 OS Viscum album.
 XX
 PN EP751221-A1.
 XX

PD 02-JAN-1997.
 XX
 PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PA (MADU) MADAUUS KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX WPI, 1997-054678/06.
 DR N-PSDB; AAT70473.
 XX
 PT Nucleic acid encoding pre-pro form of mistletoe lectin - for
 PT therapeutic or diagnostic use
 XX
 PS Claim 12; Fig 4C; 30pp; German.
 XX
 CC Mistletoe lectin is a cytotoxic agent that has been used for tumour
 CC therapy. It can be used in immunotoxins and medicaments. Nucleic
 CC acid fragments can be used in diagnostic methods. Mistletoe lectin (
 CC AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).
 XX
 SQ Sequence 564 AA;
 Query Match 89.1%; Score 1269.5; DB 18; Length 564;
 Best Local Similarity 88.9%; Pred. No. 3e-121;
 Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
 QY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAORFVLELTNOGDSITAA 60
 Db 43 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAORFVLELTNOGDSITAA 102
 QY 61 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLEFTGTRSSLPFTGSYTLDERYAGHRDQIP 120
 Db 103 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLEFTGTRSSLPFTGSYTLDERYAGHRDQIP 162
 QY 121 LGIEELIQSVSALRYPGGSTRQAARSLILILQIMISEAARFNPIFWARQYINSGESFLPD 180
 Db 163 LGIDQLIOSYVTLRFPGGSTRQAARSLILILQIMISEAARFNPIFWARQYINSGESFLPD 222
 QY 223 VYMLEETSMGQOSTVOVQOSTDGVFNPFRLGISTGNFVTLISNVBDVYIASLAIMLFCRD 282
 QY 241 RPSSSDVRYWPLYIRPVLENLSCAVDDYCTASEPTVRIY 279
 Db 283 RPSSSEVRYWPLYIRPVYI-----ADVTCSASEPTVRIY 316
 RESULT 5
 AAM90127
 ID AAM90127 standard; Protein; 564 AA.
 XX
 AC AAM90127;
 XX
 DT 30-APR-1999 (first entry)
 XX
 DE Mistletoe lectin prepro-protein.
 XX
 KW ML; mistletoe; lectin; ML; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer.
 XX
 OS Viscum album.
 XX
 PN EP84388-A1.
 XX
 PD 16-DEC-1998.
 XX
 PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX

PR 26-JUN-1995; 98EP-0105660.
XX
XX (MADU) MADAU KOELN AG.
XX
PI Baur A, Eck J, Lentzen H, Zinke H;
XX WPI: 1999-026582/03.
DR N-PSDB; AAV74182.
XX
PT New transgenic plant expressing mistletoe lectin - useful for
PT producing recombinant lectin in e.g. cancer diagnosis and therapy
XX
PS Claim 1a; Fig 4c; 30pp; German.
XX
CC This invention describes a novel transgenic plant transformed with a
CC vector capable of encoding a mistletoe (Viscum album) lectin
CC preproprotein or a biologically active fragment. The specification
CC also describes a polypeptide produced by a plant where the polypeptide
CC exhibits at least one enzymatic modification other than the glycosylation
CC that occurs in Viscum album or the polypeptide is a fusion protein, a
CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
CC polypeptide or the polypeptide dimer. The plants are used for large-scale
CC production of mistletoe lectin for diagnostic or therapeutic purposes
CC (e.g. in cancer therapy). This sequence represents the mistletoe lectin
CC used in the method of the invention.
XX
SQ Sequence 564 AA;

Query Match 89.1%; Score 1269.5; DB 20; Length 564;
Best Local Similarity 88.9%; Pred. No. 3e-121;
Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;

QY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLLKSTIPVSDAORFVLELTNOGSDSTAA 60
DB 43 HQTGDEYFRFTLLRDYVSSGFSNEIPLLKSTIPVSDAORFVLELTNOGSDSTAA 102
QY 61 IDVTNLYVVAQAGDQSYFRLDAPDGAERHLFTGTTSSRPFQSGYDLRRYAGHRQIP 120
DB 103 IDVTNLYVVAQAGDQSYFRLDAPDGAERHLFTGTTSSRPFQSGYDLRRYAGHRQIP 162
QY 121 LGIELIQSVASALRYPGSGSTRAQARSLILLIOMISEAARENPJFWRARQYINSGESFLPD 180
DB 163 LGIDQLIOSVYALRFPGSGSTRTQARSILLIOMISEAARENPJLMRRARQYINSGASFLPD 222
QY 181 MYMLEETSMGQOSTQYQOSTDGVFNPFRLGISTGNFVLSNVDYIASLAIMLFVCRD 240
DB 223 VYMLEETSMGQOSTQYQOSTDGVFNPFRLGISTGNFVLSNVDYIASLAIMLFVCRD 282
QY 241 RPSSSDVRYWPIVIRPVLENSGAVDDVTCASEPTVRIY 279
DB 283 RPSSSDVRYWPIVIRPVYI-----ADVTCASEPTVRIY 316

RESULT 6
AAV25970
ID AAV25970 standard; protein; 533 AA.
XX
AC AAV25970;
XX
DT 18-OCT-1999 (first entry)
XX
DE Mistletoe lectin protein consensus sequence 1.
XX
KW Mistletoe; lectin; antitumor; immunostimulant; A-chain; M.L.A. immunity;
KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW cancer; cytotoxicity; antigen; isoform.
XX
OS Viscum album.
XX
XX
FH Key Location/Qualifiers
FH Misc-difference 15
FT /label= Asp, Glu
FT

FT Misc-difference 63
FT /label= Gly, Glu
FT Misc-difference 66
FT /label= Ile, Val
FT Misc-difference 75
FT /label= Leu, Ala
FT Misc-difference 107
FT /label= Asp, Arg or none
FT Misc-difference 113
FT /label= Asn, Thr
FT Misc-difference 117
FT /label= Pro, Thr
FT Misc-difference 134
FT /label= Asp, Glu
FT Misc-difference 141
FT /label= Ser, Thr
FT Misc-difference 145
FT /label= Phe, Tyr
FT Misc-difference 152
FT /label= Thr, Ala
FT Misc-difference 177
FT /label= Ala, Tyr
FT Misc-difference 180
FT /label= Tyr, Asp
FT Misc-difference 185
FT /label= Ala, Glu
FT Misc-difference 191
FT /label= Val, Met
FT Misc-difference 219
FT /label= Ile, Phe
FT Misc-difference 224
FT /label= Pro, Ser
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
FT Misc-difference 287
FT /label= Asn, Ser
FT Misc-difference 290
FT /label= Cys, Arg
FT Misc-difference 325
FT /label= Gly, Asn
FT Misc-difference 364
FT /label= Gly, Asp
FT Misc-difference 426
FT /label= Gly, Gln
FT Misc-difference 435
FT /label= Val, Asp
FT Misc-difference 439
FT /label= Gln, Lys
FT Misc-difference 442
FT /label= Gly or none
FT Misc-difference 443
FT /label= Arg, Lys
FT Misc-difference 464
FT /label= Cys, Ser, Val
FT Misc-difference 480
FT /label= Ala, Gly
FT Misc-difference 481
FT /label= Gly, Ala
FT Misc-difference 483
FT /label= Ser, Gly
FT Misc-difference 484
FT /label= Gly, Ser
FT Misc-difference 493
FT /label= Gly, Tyr
FT Misc-difference 500
FT /label= Asn, Ser, Thr, Lys
FT Misc-difference 501
FT /label= Ser, Gly
FT Misc-difference 502

FT	Misc-difference	503	/label= Leu, Pro
FT	Misc-difference	504	/label= Ala, Met
FT	Misc-difference	504	/label= Met, Val
FT	Misc-difference	533	/label= Pro, Phe
XX			
PN	DE19804210-A1.		
PD	12-AUG-1999.		
PX			
PF	03-FEB-1998:	98DE-1004210.	
PR	03-FEB-1998:	98DE-1004210.	
PA	(BIOSYN ARZNEIMITTEL GMBH.		
PI	Morris P, Stiefel T, Voelter W,	Welters P;	
DR	WPI; 1999-445335/38.		
PT	Preparation of mistletoe lectins in heterologous systems,		
PT	particularly for use as anticancer agents and immunostimulants		
PS	Claim 1; Page 25-26; 78pp; German.		
CC	This invention describes a novel mistletoe lectin (I) and its fragments		
CC	which have antitumour and immunostimulatory activity. The A-chain (MLA)		
CC	of the mistletoe lectin binds to, and inactivates, the 28S subunit of		
CC	ribosomes. Non-cytotoxic forms of (I) activate T-cell and		
CC	lymphokine-producing macrophages, so stimulate immunity. (I) and its		
CC	fragments are used to treat uncontrolled cell growth (particularly of the		
CC	cancers) and if they lack cytotoxicity, to increase the strength of the		
CC	immune response, particularly to a co-administered antigen		
CC	(tumour-associated, bacterial or viral). The method allows production of		
CC	mistletoe lectin, and its individual chains, in many different isoforms		
CC	and on a large scale, at any time of the year. Recombinant products are		
CC	free from toxins present in natural mistletoe extracts. This sequence		
CC	represents a consensus sequence of the mistletoe lectin described in the		
CC	specification.		
XX			
SQ	Sequence	533 AA;	
Query Match	85.0%;	Score 1211;	DB 20; Length 533;
Best Local Similarity	86.4%;	Pred. 2.8e-115;	
Matches 242;	Conservative	6;	Mismatches 26; Indels 6; Gaps 2
QY	1 HQTGTGEFFRITLLRDVYSSGFSFSENIPLLRQSTIPVSDAQREVLVELTNOGGDSITAA 60		
DY			
DY	10 HQTTGXEFREFITLLRDVYSSGSFSFSENIPLLRQSTIPVSDAQREVLVELTNOGGDSYTAA 69		
QY	61 IDVTNLYVAAYQAGDGSYFLRDAPDGAERHLFTGTR--SLPFGGSTTDLERTAAGRHDQI 119		
DY			
DY	70 IDVTNNXYVAAYQGDDSYFLRDAPGAEHTLFTGTRRSSLPFGGSYXDLERTAVAGHRDQI 129		
QY	120 PLGIETELIVSAALRRPGSGSTRQAQRSLILILQMISEARFNPIFMWAROYINSGESFEP 179		
DY			
DY	130 PLGITXQLIGSVALKRPKPGSGSTRQAQRSIIILIQMISEARFNPILMRKRXQINSGSFEP 189		
QY	180 DMYMLEETSWMGOOSTOVQOOSTDVGFVNNEFRILGISTGNFVTLISNVRDVIASLAIMLFVCR 239		
DY			
DY	190 DYXVLELETSMWGQOSTOVQOOSTDVGFVNNEFRILGISTGNFVTLISNVRDVIASLAIMLFVCG 249		
QY	240 DRPSSSPVRKWPVLTIRPVLENESCAVDVDTCTASEPTVAIV 279		
DY			
DY	250 ERPSSSDVRKWPVLTIRPVI-----ADVVCSASEPTVAIV 284		
RESULT 7			
AAY25973			
ID	AAY25973 standard; protein: 533 AA.		
XX			

AC	AAV25973;	
XX	18-Oct-1999	(first entry)
XX	Mistletoe lectin protein consensus sequence 2.	
DE		
XX	Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLN; immunity;	
KW	rIibozyme 285 subunit; non-cytotoxic; T-cell activation; immune response;	
KW	lymphokine-producing macrophage; uncontrolled cell growth; treatment;	
KW	cancer; cytotoxicity; antigen; isoform.	
XX		
OS	Viscum album.	
XX		
FH	Key	Location/Qualifiers
FT	Misc-difference	15
FT		/label= Asp, Glu
FT	Misc-difference	63
FT		/label= Gly, Gln
FT	Misc-difference	66
FT		/label= Ile, Val
FT	Misc-difference	75
FT		/label= Leu, Ala
FT	Misc-difference	107
FT		/label= Asp, Arg, none
FT	Misc-difference	113
FT		/label= Asn, Thr
FT	Misc-difference	117
FT		/label= Pro, Thr
FT	Misc-difference	134
FT		/label= Asp, Glu
FT	Misc-difference	141
FT		/label= Ser, Thr
FT	Misc-difference	145
FT		/label= Phe, Tyr
FT	Misc-difference	152
FT		/label= Thr, Ala
FT	Misc-difference	177
FT		/label= Ala, Tyr
FT	Misc-difference	180
FT		/label= Tyr, Asp
FT	Misc-difference	185
FT		/label= Ala, Glu
FT	Misc-difference	191
FT		/label= Val, Met
FT	Misc-difference	219
FT		/label= Ile, Phe
FT	Misc-difference	224
FT		/label= Phe, Ser
FT	Misc-difference	225
FT		/label= Pro, Thr
FT	Misc-difference	232
FT		/label= Thr, Ser
FT	Misc-difference	236
FT		/label= Asp, Ser
FT	Misc-difference	287
FT		/label= Asn, Ser
FT	Misc-difference	290
FT		/label= Cys, Arg
FT	Misc-difference	325
FT		/label= Gly, Asn
FT	Misc-difference	364
FT		/label= Gly, Asp
FT	Misc-difference	427
FT		/label= Gly, Gln
FT	Misc-difference	435
FT		/label= Val, Asp
FT	Misc-difference	439
FT		/label= Gln, Lys
FT	Misc-difference	442
FT		/label= Gly, none
FT	Misc-difference	443
FT		/label= Arg, Lys
FT	Misc-difference	464

FT /label- Cys, Arg
 FT Misc-difference 325
 FT /label- Gly, Asn
 FT Misc-difference 364
 FT /label- Gly, Asp
 FT Misc-difference 426
 FT /label- Gly, Gln
 FT Misc-difference 435
 FT /label- Val, Asp
 FT Misc-difference 439
 FT /label- Gln, Lys
 FT Misc-difference 442
 FT /label- Gly, none
 FT Misc-difference 443
 FT /label- Arg, Lys
 FT Misc-difference 464
 FT /label- Cys, Ser, Val
 FT Misc-difference 480
 FT /label- Ala, Gly
 FT Misc-difference 481
 FT /label- Gly, Ala
 FT Misc-difference 483
 FT /label- Ser, Gly
 FT Misc-difference 484
 FT /label- Gly, Ser
 FT Misc-difference 493
 FT /label- Gly, Tyr
 FT Misc-difference 500
 FT /label- Asn, Ser, Thr, Lys
 FT Misc-difference 501
 FT /label- Ser, Gly
 FT Misc-difference 502
 FT /label- Leu, Pro
 FT Misc-difference 503
 FT /label- Ala, Met
 FT Misc-difference 504
 FT /label- Met, Val
 FT Misc-difference 533
 FT /label- Pro, Phe
 DE19804210-A1.
 12-AUG-1999.
 03-FEB-1998: 98DE-1004210.
 03-FEB-1998: 98DE-1004210.
 (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 PI Morris P, Stiefel T, Voelter W, Welters P;
 DR WPI: 1999-445335/38.
 PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Claim 40: Page 37-38; 78pp; German.
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of the mistletoe lectin described in the
 CC specification.

XX
 SO Sequence 533 AA;
 Query Match 85.0%; Score 1211; DB 20; Length 533;
 Best Local Similarity 86.4%; Pred. No. 2,8e-115;
 Matches 242; Conservative 6; Mismatches 26; Indels 6; Gaps 2;
 QY 1 HOTGDEYFRITLLRDVSSGSPSNEPIPLROSTIVSDORFVLELNOGDSITAA 60
 DB 10 HOTGEXEYFRITLLRDVSSGSPSNEPIPLROSTIVSDORFVLELNOGDSITAA 69
 QY 61 IDVTNLYVVAQAGDQSYFLRDAPGAEHLFTGTR-SSLPEFGSYTDLERYAGHRDQI 119
 DB 70 IDVTNXXVVAQAGDQSYFLRDAPGAEHLFTGTRXSSLPFGGSYXDLERYAGHRDQI 129
 QY 120 PLGIEELIQSVSALRYPGGSTRQAQARSLIILQMISEARFNPIFMRARQYINSGESFLP 179
 DB 130 PLGIXOLIQSVXALRXPGGSTRQAQARSLIILQMISEARFNPILMRXRXQYINSGSXFLP 189
 QY 180 DMYMLEFTSMGOOSTOVQSTDGVPNNPRLIGSTGFEVTLNVRDVIASLAIMLFCVCR 239
 DB 190 DYMMELETSWGOOSTOVQSHSTDGVFNNPRLAIXXGFEVTLXNVXVYIASLAIMLFCVCG 249
 QY 240 DRPSSDVRYWPLYRPLYLENSGAVDDVTCYASEPTVRIIV 279
 DB 250 ERPSSDVRYWPLYRPLYI-----ADDVTCASEPTVRIIV 284
 RESULT 9
 AAB47090
 ID AAB47090 standard; Protein: 254 AA.
 XX
 AC AAB47090;
 DT 16-MAY-2001 (first entry)
 DE
 XX
 DE A-chain isoform for biosynthesis of a Korean mistletoe lectin #1.
 KW Isoform: A-chain; B-chain: biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KMHBP; KML-110; KML-11L;
 KW heparin binding protein.
 XX
 OS Viscum album coloratum.
 XX
 PN EP1074560-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 27-JUL-2000; 2000EP-0402160.
 XX
 PR 27-JUL-1999; 99KR-0030638.
 XX
 PA (MIST-) MISTLE BIOTECH CO LTD.
 XX
 PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PT Kang T, Park C;
 XX
 DR WPI: 2001-171044/18.
 XX
 DR N-PSDB; AAC85472.
 PT Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 XX
 PS Claim 2: Page 25-26; 62pp; English.
 CC The sequences given in AAB47090-92 are isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumours. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-MJ25 lymphoma.

SQ Sequence 254 AA;
 Query Match 81.1%; Score 1155; DB 22; Length 254;
 Best Local Similarity 93.1%; Pred. No. 5.3e-110;
 Matches 228; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSNFIPLRQSTIPVSDAORFVLELINOAGDSITAA 60
 DB 10 HQTGDEYFRFTLLRDYVSSGFSNFIPLRQSTIPVSDAORFVLELINOAGDSITAA 69
 QY 61 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLEFTGTTSSSLPFTGSYTDLERVAGHRDQ 120
 DB 70 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLEFTGTTSSSLPFTGSYTDLERVAGHRDQ 129
 QY 121 LGIEELIOSVSAALRYPGSGSTRPAORSLILILOMISEARPNPIFWRARQYINGESFLP 180
 DB 130 LGREELIOSVSAALRYPGSGSTRPAORSLILILOMISEARPNPILMRRQYISSGGSFLP 189
 QY 181 MYMLEETSMWGQOSTQVQOSTDGVFNPFRLGISTGNFVTLISNVRDYIASLAIMLFVC 240
 DB 190 TYILQLETSWGQOSTQVQOSTDGVFNPNPRLITISTGVFVTLISNVRDYIASLAIMLFVC 249
 QY 241 RPSSS 245
 DB 250 RPSSS 254

RESULT 10
 AAY25981
 ID AAY25981 standard; Protein; 256 AA.
 AC AAY25981;
 DT 18-OCT-1999 (first entry)
 XX Mistletoe lectin A2 protein fragment.
 DE
 XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A2.
 XX
 OS Viscum album.
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI, 1999-445335/38.
 DR N-PSDB; AA209105.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Claim 8; Fig 3B; 78pp; German.
 CC
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of

CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of the mistletoe lectin A2 protein.
 CC
 SQ Sequence 256 AA;
 Query Match 80.8%; Score 1152; DB 20; Length 256;
 Best Local Similarity 93.1%; Pred. No. 1.1e-109;
 Matches 230; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 HQTGDEYFRFTLLRDYVSSGFSNFIPLRQSTIPVSDAORFVLELINOAGDSITAA 60
 DB 10 HQTGDEYFRFTLLRDYVSSGFSNFIPLRQSTIPVSDAORFVLELINOAGDSITAA 69
 QY 61 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLEFTGTT--RSSLPFTGSYTDLERVAGHRDQ 118
 DB 70 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLEFTGTTTRRSSLPFTGSYTDLERVAGHRDQ 129
 QY 119 IPLGIELIOSVSAALRYPGSGSTRPAORSLILILOMISEARPNPIFWRARQYINGESFL 178
 DB 130 IPLGIELIOSVSAALRYPGSGSTRPAORSLILILOMISEARPNPILMRRQYISSGGSFL 189
 QY 179 PDMYMLEETSMWGQOSTQVQOSTDGVFNPFRLGISTGNFVTLISNVRDYIASLAIMLFVC 238
 DB 190 PDMYMLEETSMWGQOSTQVQOSTDGVFNPNPRLAISTGNFVTLISNVRDYIASLAIMLFVC 249
 QY 239 RDRPSSS 245
 DB 250 GERPSSS 256

RESULT 11
 AAY25984
 ID AAY25984 standard; Protein; 256 AA.
 AC AAY25984;
 DT 18-OCT-1999 (first entry)
 XX Mistletoe lectin A2 (variant) protein fragment.
 DE
 XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A2.
 XX
 OS Viscum album.
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI, 1999-445335/38.
 DR N-PSDB; AA209108.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Disclosure; Fig 6B; 78pp; German.
 CC
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and

CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A2 protein variant.

XX Sequence 256 AA:

Query Match 80.8%; Score 1152; DB 20; Length 256;

Best Local Similarity 93.1%; Pred. No. 1.1e-109;

Matches 220; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 HQTGGDEYFRFTLLRDYSSGFSNEIPLLRQSTIPVSDAQRFVLELTNGGDSITAA 60

DB 10 HQTGGDEYFRFTLLRDYSSGFSNEIPLLRQSTIPVSDAQRFVLELTNGGDSITAA 69

QY 61 IDVTNLYVVAAYQAGDOSYFLRDAPDGAERHLEFTGTT--RSLEPFGSYTDLERYAGHRDQ 118

DB 70 IDVTNLYVVAAYQAGDOSYFLRDAPDGAERHLEFTGTTDRSSLPFGSYTDLERYAGHRDQ 129

QY 119 IPIGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPIFMRARQYINGSGESFL 178

DB 130 IPIGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPIFMRARQYINGSGESFL 189

QY 179 PDWYMLELFTSMGQOSTQVOQSTDGVFNPNPRLGISTGNFTVLTNSNRDVYASLAIMLFVC 238

DB 190 PDWYMLELFTSMGQOSTQVOHSTDGVFNPNPRLAIPGNFVLTNVRVYASLAIMLFVC 249

QY 239 RDRPSSS 245

DB 250 GERPSSS 256

RESULT 12

AAW64661 ID AAW64661 standard; Protein; 252 AA.

XX AAW64661;

DT 23-OCT-1998 (first entry)

XX Mistletoe rMLA variant protein.

XX Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;

KW intracellular; processing module; protease recognition; targeting module;

KM internalisation; treatment; disorder; cell proliferation; activation;

XX autoimmune disease; allergy; tumour; ricin; translocation; ss.

XX Viscum album.

XX Key Location/Qualifiers

FT Protein 1..252

XX /note="partial"

XX W09829540-A2.

XX 09-JUL-1998.

XX 02-JAN-1998; 98WO-EP00009.

XX 02-JAN-1997; 97EP-0100012.

XX (BRA1-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.

XX Eck J, Schmidt A, Zinke H;

XX WPI: 1998-388122/33.

XX DR N-PSDB: AAV51343.

PT Nucleic acid encoding fusion protein containing mistletoe lectin A
 PI chain - useful for treatment of proliferative and autoimmune
 PS diseases, allergies and tumours

XX Disclosure; Fig 11a'; 115pp; German.

CC This sequence encodes a variant mistletoe lectin A-chain, rMLA. This
 CC sequence can be used in the construction of a fusion protein which
 CC comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC e.g. by injection or topically but especially by intravenous injection,
 CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.

XX Sequence 252 AA:

Query Match 79.4%; Score 1131; DB 19; Length 252;

Best Local Similarity 90.5%; Pred. No. 1.5e-107;

Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGGDEYFRFTLLRDYSSGFSNEIPLLRQSTIPVSDAQRFVLELTNGGDSITAA 60

DB 10 HQTGGDEYFRFTLLRDYSSGFSNEIPLLRQSTIPVSDAQRFVLELTNGGDSITAA 69

QY 61 IDVTNLYVVAAYQAGDOSYFLRDAPDGAERHLEFTGTTSSLPFGSYTDLERYAGHRDQIP 120

DB 70 IDVTNLYVVAAYQAGDOSYFLRDAPDGAERHLEFTGTTSSLPFGSYTDLERYAGHRDQIP 129

QY 121 LGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPIFMRARQYINGSGESFLPD 180

DB 130 LGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPIFMRARQYINGSGESFLPD 189

QY 181 MYWYMLELFTSMGQOSTQVOQSTDGVFNPNPRLGISTGNFTVLTNSNRDVYASLAIMLFVC 240

DB 190 MYWYMLELFTSMGQOSTQVOHSTDGVFNPNPRLAIPGNFVLTNVRVYASLAIMLFVC 249

QY 241 RPS 243

DB 250 RPS 252

RESULT 13

AAW10022 ID AAW10022 standard; Protein; 253 AA.

XX AAW10022;

XX 18-DEC-1997 (first entry)

XX Prepro mistletoe lectin A chain.

XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.

XX Viscum album.

XX EP751221-A1.

XX 02-JAN-1997.

XX

PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PA (MADU) MADDAUS KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX
 DR WPI; 1997-054678/06.
 DR N-PSDB; AAT70474.
 XX
 PT Nucleic acid encoding pre-pro form of mistletoe lectin - for
 PT therapeutic or diagnostic use
 XX
 PS Claim 12; Fig 4A; 30pp; German.
 CC Mistletoe lectin is a cytotoxic agent that has been used for tumour
 CC therapy. It can be used in immunotoxins and medicaments. Nucleic
 CC acid fragments can be used in diagnostic methods. Mistletoe lectin (
 CC AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).
 CC
 SO Sequence 253 AA;

Query Match 79.4%; Score 1131; DB 18; Length 253;
 Best Local Similarity 90.5%; Pred. No. 1.5e-107;
 Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFNEIPLRQSTIPVSDAQRVLYVELTNGGDSITAA 60
 |||||
 DB 11 HQTGEYFRFTLLRDYVSSGFSFNEIPLRQSTIPVSDAQRVLYVELTNGGDSITAA 70
 QY 61 IDVTNLYVVAQAGDSYFLRDAPDGAERHLEFTGTRSSLPFGSYTDLERVAGHRDQIP 120
 |||||
 DB 71 IDVTNLYVVAQAGDSYFLRDAPDGAERHLEFTGTRSSLPFGSYTDLERVAGHRDQIP 130
 QY 121 LGIEELIQSVSLARYPGGSTRQAARSLIILMISEARFNPIFWRARQYINSGSFLPD 180
 |||||
 DB 131 LGIDQLIQSVTLARFPGGSTRQAARSLIILMISEARFNPIFWRARQYINSGSFLPD 190
 QY 181 MYMLELTSWGQOSTVOVQSTGDFVNNPRLIGISTGNFVTLNVRDYIASLAIMLFVCRD 240
 :|||||
 DB 191 MYMLELTSWGQOSTVOVQSTGDFVNNPRLIPGNFVTLNVRDYIASLAIMLFVCGE 250
 QY 241 RPS 243
 |||
 DB 251 RPS 253

RESULT 14
 AAM60125
 ID AAM90125 standard; Protein; 253 AA.
 XX

AC AAM90125;
 XX
 DT 30-APR-1999 (first entry)
 XX

DE Mistletoe ML A-chain protein.
 XX

KW MI; mistletoe; lectin; MLA; A-chain; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer.
 XX

OS Viscum album.
 XX

PN EP864388-A1.
 XX

PD 16-DEC-1998.
 XX

PF 26-JUN-1995; 95EP-0109949.
 XX

PR 26-JUN-1995; 95EP-0109949.
 PR 26-JUN-1995; 98EP-0105660.
 XX

PA (MADU) MADDAUS KOELN AG.
 XX

PI Baur A, Eck J, Lentzen H, Zinke H;
 XX

DR WPI; 1999-026582/03.
 DR N-PSDB; AAV74180.
 XX

PT New transgenic plant expressing mistletoe lectin - useful for
 PT producing recombinant lectin in e.g. cancer diagnosis and therapy
 XX

PS Disclosure; Fig 4a; 30pp; German.
 XX

CC This invention describes a novel transgenic plant transformed with a
 CC vector capable of encoding a mistletoe (Viscum album) lectin
 CC preproprotein or a biologically active fragment. The specification
 CC also describes a polypeptide produced by a plant where the polypeptide
 CC exhibits at least one enzymatic modification other than the glycosylation
 CC that occurs in Viscum album or the polypeptide is a fusion protein, a
 CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 CC polypeptide or the polypeptide dimer. The plants are used for large-scale
 CC production of mistletoe lectin for diagnostic or therapeutic purposes
 CC (e.g. in cancer therapy). This sequence represents the mistletoe lectin
 CC A-chain which is contained in expression vector pT7MLA.
 CC
 SO Sequence 253 AA;

Query Match 79.4%; Score 1131; DB 20; Length 253;
 Best Local Similarity 90.5%; Pred. No. 1.5e-107;
 Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFNEIPLRQSTIPVSDAQRVLYVELTNGGDSITAA 60
 |||||
 DB 11 HQTGEYFRFTLLRDYVSSGFSFNEIPLRQSTIPVSDAQRVLYVELTNGGDSITAA 70
 QY 61 IDVTNLYVVAQAGDSYFLRDAPDGAERHLEFTGTRSSLPFGSYTDLERVAGHRDQIP 120
 |||||
 DB 71 IDVTNLYVVAQAGDSYFLRDAPDGAERHLEFTGTRSSLPFGSYTDLERVAGHRDQIP 130
 QY 121 LGIEELIQSVSLARYPGGSTRQAARSLIILMISEARFNPIFWRARQYINSGSFLPD 180
 |||||
 DB 131 LGIDQLIQSVTLARFPGGSTRQAARSLIILMISEARFNPIFWRARQYINSGSFLPD 190
 QY 181 MYMLELTSWGQOSTVOVQSTGDFVNNPRLIGISTGNFVTLNVRDYIASLAIMLFVCRD 240
 :|||||
 DB 191 MYMLELTSWGQOSTVOVQSTGDFVNNPRLIPGNFVTLNVRDYIASLAIMLFVCGE 250
 QY 241 RPS 243
 |||
 DB 251 RPS 253

RESULT 15
 AAM64659
 ID AAM64659 standard; Protein; 252 AA.
 XX

AC AAM64659;
 XX

DT 23-OCT-1998 (first entry)
 XX

DE Mistletoe rMLA protein.
 XX

KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation.
 XX

OS Viscum album.
 XX

PN WO9829540-A2.
 XX

PD 09-JUL-1998.
 XX

PF 02-JAN-1998; 98WO-EP00009.
 XX

```
XX 02-JAN-1997; 97EP-0100012.
PR
XX (BRAI-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
PA
XX Eck J, Schmidt A, Zinke H;
PI
XX MPI; 1998-388122/33.
DR
XX N-PSDB; AAV51341.
XX
PT Nucleic acid encoding fusion protein containing mistletoe lectin A
PT chain - useful for treatment of proliferative and autoimmune
PT diseases, allergies and tumours
XX
PS Disclosure: Fig 11a; 115pp; German.
XX
CC This sequence represents a lectin A-chain, rMLA, isolated from mistletoe.
CC This sequence can be used in the construction of a fusion protein which
CC comprises an effector module that is cytotoxic intracellularly, a
CC processing module covalently bonded to the effector module and
CC containing a protease recognition sequence, and a targeting module
CC covalently bonded to the processing module, able to bind specifically to
CC the surface of a cell so as to mediate internalisation of the fusion
CC protein. Such a fusion protein can be used for treating disorders
CC involving proliferation and/or elevated activation of cells, especially
CC autoimmune disease, allergy and tumours. The proteins can be administered
CC e.g. by injection or topically but especially by intravenous injection,
CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.
CC Fusion proteins can develop toxic activity in a wide range of target
CC cells. The processing module prevents extracellular dissociation, and
CC fusion proteins based on mistletoe lectin A-chain are far more active
CC than those based on ricin and do have the associated problems of
CC non-specific toxicity. The protein may be expressed in a non-glycosylated
CC form that does not bind to sugar receptors in the liver, and which has a
CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
CC it actively assists in translocation of the ML A-chain from the
CC endoplasmic reticulum to the cytoplasm.
XX
SQ Sequence 252 AA;

```

Query Match 79.1%; Score 1127; DB 19; Length 252;
Best Local Similarity 90.5%; Pred. No. 3.9e-107;
Matches 219; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```
QY 1 HQTGDEFFRFTLLRDVYSSSEFSENEPLRQSTIPYSDAQRFLVELTNOGDSITAA 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 11 HQTGEFFRFTLLRDVYSSSEFSENEPLRQSTIPYSDAQRFLVELTNOGDSITAA 70
QY 61 IDVTNLYVAVYAGDQSYFLRDAPGAEERHLEFTGTTRSSLPPTGSYTDLERYAGHRDIP 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 71 IDVTNLYVAVYAGDQSYFLRDAPGAEERHLEFTGTTRSSLPPTGSYTDLERYAGHRDIP 130
QY 121 LGIEELIQSVSLARYPGGSTRAQARSLLILQMISEARFNPIFWARQYINSGESFLPD 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 131 LGIDQLIQSVTLRFPGGSTRTQARSLLILQMISEARFNPIFLMARQYINSGESFLPD 190
QY 181 MYMLELFTSMGQSTQVQOQSTDGVANNPFRIGISTGNFVTLSNVADVTASLAIMLFVCRD 240
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 191 VYMLELFTSMGQSTQVQOQSTDGVANNPFRILAIPPGNFVTLTNVADVTLASLAIMLFVCGE 250
QY 241 RP 242
   ||
DB 251 RP 252

```

Search completed: March 18, 2003, 08:25:13
Job time : 52.0704 secs


```

? RESULT 3
? US-08-776-059-39
? Sequence 39, Application US/08776059B
? Patent No. 6271368
? GENERAL INFORMATION:
? APPLICANT: LENTZEN, Hans
? APPLICANT: ECK, Jürgen
? APPLICANT: BAUR, Axel
? APPLICANT: ZINKE, Holger
? TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
? FILE REFERENCE: 674503-2003
? CURRENT APPLICATION NUMBER: US/08/776,059B
? CURRENT FILING DATE: 1999-06-19
? EARLIER APPLICATION NUMBER: PCT/EP96/02273
? EARLIER FILING DATE: 1996-06-25
? EARLIER APPLICATION NUMBER: 95109949.8
? EARLIER FILING DATE: 1995-06-26
? NUMBER OF SEQ ID NOS: 56
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 39
? LENGTH: 235
? TYPE: PRP
? ORGANISM: Viscum album

```

Dh 13 TSOSYKOFTEALREEL-RGGLIHDPVLPDPT-TLOERNRYITVELSNSDTESEIEVGIDV 70

	Query Match	32.7%	Score 466.5;	DB 1;	Length 250;
	Best Local Similarity	43.7%	Pred. NO. 1.3e-41;		
	Matches 108;	Conservative 36;	Mismatches 78;	Indels 25;	Gaps 6;
Qy	4 TGDYEFITLLRDYVSSGSFNSNELLPLLRÖSTIPVSQAQRFVLVELLNOGGDSTITAAIDV	63			
	: :	:	: :	:	: :
Dd	13 TSQSKPCLEAIRRETL-RGGLHDIHPVDPP-TLOERNRRTIVSLSDNESIEVGDIW	70			

OY	64	TNLYVVAQAQDSDSYFLRBDAPDGAERHLFTGTRRSLSLPEFGSTYDLERYAGH-RDQIPICG	122
Dd	71	TMAVVAARACTOSYLFRDAPSSASDYLFETGDOSHLPFYGTGDDLERMHQSROOPIPLG	130
OY	123	IEELIOSVALRYPGSGSTRAOARSLITLIOMISEAFENFIEMRAROYINSGESFLPDWM	182
Dd	131	LQALTHHGSIFFRRSCGNONEEKARLLIYIOWVAEAAEFRIISNRVKYSIOTGTAFODDA	190
OY	183	MLELETSMWGQOSTIOVGQOSTDGVEFNNEPRIGISTGNFVTLSNVHD-----VIASLT	231
Dd	191	MISLENNMW-DNLRGVQESVDTFPNO-----VTLININNEPYIVDSLSHPYAVVL	239
OY	232	AIMLEVC	238
Dd	240	AIMLEVC	246

```

1      RESULT 5
2      US-08-485-286-71
3      : Sequence 71, Application US/08485286
4      : Patent No. 5646026
5      : Patent No. 5646026 5646119
6      :
7      : GENERAL INFORMATION:
8      : APPLICANT: WALSH, TERENCE A
9      : APPLICANT: HEY, TIMOTHY D
10     : APPLICANT: MORGAN, ALICE ER
11     : TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEIN, INACTIVE
12     : TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
13     : TITLE OF INVENTION: USING
14     : NUMBER OF SEQUENCES: 81
15     : CORRESPONDENCE ADDRESS:
16     : ADDRESSEE: ANDREA T. BORUCKI
17     : STREET: 9330 ZIONSVILLE ROAD
18     : CITY: INDIANAPOLIS
19     : STATE: IN
20     : COUNTRY: US
21     : ZIP: 46268
22     :
23     : COMPUTER READABLE FORM:
24     : MEDIUM TYPE: Floppy disk
25     : COMPUTER: IBM PC compatible
26     : OPERATING SYSTEM: PC-DOS/MS-DOS
27     : SOFTWARE: PatentIn Release #1.0, Version #1.25
28     : CURRENT APPLICATION DATA:
29     : APPLICATION NUMBER: US/08/485,286
30     : FILING DATE:
31     : CLASSIFICATION: 435
32     : PRIOR APPLICATION DATA:
33     : APPLICATION NUMBER: US 08/378761
34     : FILING DATE: 26-JAN-1995
35     : ATTORNEY/AGENT INFORMATION:
36     : NAME: BORUCKI, ANDREA T
37     : REGISTRATION NUMBER: 33651
38     : REFERENCE/DOCKET NUMBER: 38272B
39     : TELECOMMUNICATION INFORMATION:
40     : TELEPHONE: (317) 337-4846
41     : INFORMATION FOR SEQ ID NO: 71:
42     : SEQUENCE CHARACTERISTICS:
43     : LENGTH: 250 amino acids
44     : TYPE: amino acid
45     : STRANDEDNESS: single
46     : TOPOLOGY: linear
47     : MOLECULE TYPE: protein
48     :
49     :
50     :
51     :
52     :
53     :
54     :
55     :
56     :
57     :
58     :
59     :
60     :
61     :
62     :
63     :
64     :
65     :
66     :
67     :
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :
625    :
626    :
627    :
628    :
629    :
630    :
6
```

```

Query Match 32.7% Score 466.5; DB 1, Length 250;
Best Match Similarity 43.7%; Pred. NO. 1.3e-81;
Matches 108; Conservative 36; Mismatches 78; Indels 25; Gaps 6;

QY 4 TGEYRFTLLRDVYSGSFSNEIPLIKOSTIPVSDAQRFLVELTNOGDSITAIADV 63
      |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 13 TSOQVQFEAKRERIKGLIHDPIDPTFLQERNRIYVELNSNSTEISLEGVIGIV 70

```

```

QY 64 TULYVAAVQAGQSEFLDADPADGAEARHLFTCTPSSSLPFGSYNDLBRVGH -RDOIHLG 122
   |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 71 TAAIVVATRAGIQTFLDAPSSASDYLFTGTDQHSLEFYGYDLERMAHQIPLG 130
   |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 123 IEEILQSVSALRYPGGSTRQAORSLILIQMISEAARFPIFMARQYINSGESFLPMY 182
   ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 131 LQALHNGISLFPFGSGNDMEERKRLIYLIIQVALEAARRYISNKNVRSIQGTGFOPDAA 190
   |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 183 MLEETSNKGQOSTVOQSTDGVFNPNPRLGISTGNFVTLSNVRD-----VIASL 231
   ||::||::||::||::||::||::||::||::||::||::||::||::||
Db 191 MISELENNM-DNLRGVQESVQDCTFPNQ-----VLTFRIRREPIYVDSLHPYAVL 239
   ||::||::||::||::||::||::||::||::||::||::||::||::||
QY 232 ATMLEFVC 238
   ||::|||::|||
Db 240 ALMLEFVC 246

```

```

1      RESULT 6
2      US-08-378-761A-77
3      : Sequence 77, Application US/08378761A
4      : Patent No. 5635384
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: WALSH, TERENCE A
9      : APPLICANT: HEY, TIMOTHY D
10     : APPLICANT: MORGAN, ALICE ER
11     :
12     : TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
13     : TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
14     : TITLE OF INVENTION: USING
15     : NUMBER OF SEQUENCES: 81
16     :
17     : CORRESPONDENCE ADDRESS:
18     : ADDRESSEE: ANDREA T. BORUCKI
19     : STREET: 9330 ZIONSVILLE ROAD
20     : CITY: INDIANAPOLIS
21     :
22     : STATE: IN
23     :
24     : COUNTRY: US
25     :
26     : ZIP: 46268
27     :
28     : COMPUTER READABLE FORM:
29     :
30     : MEDIUM TYPE: Floppy disk
31     : COMPUTER: IBM PC compatible
32     : OPERATING SYSTEM: PC-DOS/MS-DOS
33     : SOFTWARE: PatentIn Release #1.0, Version #1.25
34     :
35     : CURRENT APPLICATION DATA:
36     : APPLICATION NUMBER: US/08/378,761A
37     : FILING DATE: 26-JAN-1995
38     : CLASSIFICATION: 435
39     :
40     : ATTORNEY/AGENT INFORMATION:
41     : NAME: BORUCKI, ANDREA T
42     : REGISTRATION NUMBER: 33651
43     : REFERENCE/DOCKET NUMBER: 38272B
44     :
45     : TELECOMMUNICATION INFORMATION:
46     : TELEPHONE: (317) 337-4846
47     : INFORMATION FOR SEQ ID NO: 77:
48     :
49     : SEQUENCE CHARACTERISTICS:
50     : LENGTH: 540 amino acids
51     : TYPE: amino acid
52     : STRANDEDNESS: single
53     :
54     : TOPOLOGY: linear
55     :
56     : MOLECULE TYPE: protein
57     :
58     : US-08-378-761A-77

```

	29.7%;	Score 423;	DB 1;	Length 540;
Query Match:	Best Local Similarity	39.7%;	Pred. No. 1,76-36;	
Matches 114;	Conservative	45;	Mismatches 106;	Indels 22; Gaps 10.
OY	4	TGDEFFRFTLLRDVSSGS-FSNELPLL-ROSTPVSADARFVLVELTNOGGDSTAAI	61	
:	:	: :: :	:: :	:
Db	17	TVEYTFETIAASHLITTCADVRHETPLVPALPNKRGLPIS--ORFLTVLSMHAELSVTLAL	74	
OY	62	DVTNLYVVAYAOQDOSGYFLRDAPDCAE-----RHLEFTTRS-SLPTPGSYLTLEKARG	115	
:	:	:	:	:
Db	75	DVTNAYVVGCRAGNSAFYFH--PDNDGEAEATHLFIDVONSFPFAAGNGDRLEQUGLI	132	
OY	116	RDOIPGLIEELLQSVALRY---PGGSTRAQAARSLILTIOMISEAARNFDIPIRAROYIN	172	

Db 133 RENIEIGTGLEDAISALYYSTGCTQIPTLARSEFVCIOIMISEARFOYIEGEMTRIR 192
QY 173 SGESEFLPDMYMLEETSMGOOSTVOQSTGVNNPRLCISTGNFVTLNVDVIAASLA 232
Db 193 YNRKSAFDPSTVITLNSMGRSLTAIOESNOGAFASPIQLORNGSKFNVDYIILPIIA 252
QY 233 IMLFVCRDRPSSDVRYMPLVIRPLENSGAVDVTCTASEPTVRIY 279
Db 253 LMVYRCAPPSSQ----FSLIRPVVFNMA--DV-CMDEPIVRIY 292

RESULT 7

US-08-485-286-77
; Sequence 77, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-286-77

Query Match 29.7%; Score 423; DB 1; Length 540;
Best Local Similarity 39.7%; Pred. No. 1.7e-36;
Matches 114; Conservative 45; Mismatches 106; Indels 22; Gaps 10;
QY 4 TGDEYFRFTLLRDYVSSGS--FSNEIPL--RSTIPVSDAORFVLELINOOGSDITAAI 61
Db 17 TVESTNTRAVRSHLTGADVREHIEPLVPRVGLPTS--QRTIVELSNHAELSVTLAL 74
QY 62 DVNLVVAAYQAGDOSYFLRDAPDAE-----RHLEFGTTRS-SLPFTGTYTDLERYAGH 115
Db 75 DVTAAYVVGCRAGNSAYFEH--PDNOEDAEALITLFTDVQNSFFAFAGGNDRLQLAGN 132
QY 116 RDQIFLGIEELIOSVALRY---PGSSTRAQARSIIILQIMISAARFNPLFWARQYIN 172

Db 133 RENIEIGTGLEDAISALYYSTGCTQIPTLARSEFVCIOIMISEARFOYIEGEMTRIR 192
QY 173 SGESEFLPDMYMLEETSMGOOSTVOQSTGVNNPRLCISTGNFVTLNVDVIAASLA 232
Db 193 YNRKSAFDPSTVITLNSMGRSLTAIOESNOGAFASPIQLORNGSKFNVDYIILPIIA 252
QY 233 IMLFVCRDRPSSDVRYMPLVIRPLENSGAVDVTCTASEPTVRIY 279
Db 253 LMVYRCAPPSSQ----FSLIRPVVFNMA--DV-CMDEPIVRIY 292

RESULT 8

US-08-356-786-10
; Sequence 10, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-10

Query Match 28.7%; Score 408.5; DB 2; Length 534;
Best Local Similarity 38.9%; Pred. No. 5.9e-35;
Matches 112; Conservative 49; Mismatches 102; Indels 25; Gaps 11;
QY 4 TGDEYFRFTLLRDYVSSGS--FSNEIPL--RSTIPVSDAORFVLELINOOGSDITAAI 61
Db 20 TVQSYNFTRAVGRITTGADVREHIEPLVPRVGLPIN--QRTIVELSNHAELSVTLAL 77
QY 62 DVNLVVAAYQAGDOSYFLRDAPDAE-----RHLEFGT--TRSSLPFTGTYTDLERYAGH 115
Db 78 DVTAAYVVGCRAGNSAYFEH--PDNOEDAEALITLFTDVQNTYTFAGFNRYRLQLAGN 135
QY 116 -RDQIFLGIEELIOSVALRY---PGSSTRAQARSIIILQIMISAARFNPLFWARQYI 171
Db 136 LRNIEIGNPLEDAISALYYSTGCTQIPTLARSPFICIOIMISEARFOYIEGEMTRIR 195

```

OY      172  NGSFSPDLEMYLEETMSQGOOSQVOOOSTGCVNPNPFRIGISNGVNTSNVDYASL  231
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      196  RYNRSAPDPEVITLLEMSQGRSLTAIDESNCGAFASPTOLRRNGSKFVDSILPITI  255

OY      232  AIMEFVGRDRSSSDVRYTWPIYRPLVENSAGADVDTCTASEPVRIVY  279
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      256  ALMYIRKAPPPSSO----FSLIRIPVFNENA--DV---CDDPEIOLV  294

```

RESULT 9
US-07-901-707-1

```

Sequence 1 Application US/07901707
Patent No. 5376546
GENERAL INFORMATION:
APPLICANT: Bernhard, Susan L.
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Steve F.
APPLICANT: Lane, Julie A.
TITLE OF INVENTION: Materials Comprising and Methods of
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/901,707
FILING DATE: 19920619
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5376546and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27129/30910
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 984-5750
TELEFAX: (312) 984-5750
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-901-707-1
Query Match      26.9%; Score 383; DB 1; Length 267;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;
QY 4 TGDEYFFRILLRDYSSGS-FSNEIPL-RGSTIPSAQRVLVELNMGSDSTAI 61
| | | : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 TVGSYTFIRAFNRRLTTGADVRRHEIVLPFRVGPIIN-QRFIVELSHHAELSVTLAV 74
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
QY 62 DVTNLYVAAYOAGDQSYFLRADPGAE---RHLEFGT-TRSSLPTGSGTYDLERYAGH 115
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 75 DVTAAYVAGVRAGNSAFFFH--PDNDADAETHLFTDVONRYTFPAFGGYDRLEQLAGN 132
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
QY 116 -RDPIPGIEELLISVALRY---PGSGTAAQRSLILLLQMISEAARFPITWRKROYI 171
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 133 LRKEIVLGNGCLPEEATLSALYYSTGTGLPFTLARSPFIQIMSEARQYLEGEMRTET 192

```

Qy 172 NSGESEFDDYMLEETSGOOSTVOOOSTGVENNFRLGISGNEVTLISNVDYASL 231
 |||||
 Db 193 RYNRRSADPEVITLENBGRSLTAIGESNGCALASPLQLRRNGSNFSYDVYSILPII 252
 |||||
 Qy 232 ALIMFVCDRRSS 244
 |||||
 Db 253 ALMYRCAIPPS 265

RESULT 10
US-07-988-430-1

[illegible]

Db 75 DVTNAYVVGVRAGNSAYFFH--PDNQEDAETHLFTDVQNRYPFAGGNDYRLQOLAGN 132
QY 116 -RDQIPLGIEELIQSVSLRY---FGSTRAQARSLIILQIMISANPNPIFMRAROYI 171
Db 133 LRENIEIENGPLEIAISALYYSTGTQOLPTLARSPFICIQIMISANRFOYIEGEMRTRI 192
QY 172 NSGESFLPDMYMLETSMGQOSTVOVQSTGVFNNPRLGISTGNFTLSNVRVYIASL 231
Db 193 RYNRSAPDPSPVITLNSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIT 252
QY 232 AIMLFVCRDRPSS 244
Db 253 ALMYRCAPPSS 265

RESULT 11
US-08-218-303-16
; Sequence 16, Application US/08218303
; Patent No. 5547867
; GENERAL INFORMATION:
; APPLICANT: Kara, Bhupendra V.
; APPLICANT: Hockney, Robert C.
; APPLICANT: Filton, John E.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,303
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/841,533
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PAK/3893/94908/MJW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO. 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-218-303-16

Query Match 26.9%; Score 383; DB 1; Length 267;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;

QY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPL-LRSTIPVSDAQRVIVLITNOGSDSTIAI 61
Db 17 TVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPIN--QRILVELSNHAEISVTLAL 74
QY 62 DVTNLYVAVQAGDOSYFLRDAPDCAE-----RHLPFTGT-TRSSLPFTSGYTDLERYAGH 115
Db 75 DVTNLYVAVQAGDOSYFLRDAPDCAE-----RHLPFTGT-TRSSLPFTSGYTDLERYAGH 132
QY 116 -RDQIPLGIEELIQSVSLRY---FGSTRAQARSLIILQIMISANPNPIFMRAROYI 171

Db 133 LRENIEIENGPLEIAISALYYSTGTQOLPTLARSPFICIQIMISANRFOYIEGEMRTRI 192
QY 172 NSGESFLPDMYMLETSMGQOSTVOVQSTGVFNNPRLGISTGNFTLSNVRVYIASL 231
Db 193 RYNRSAPDPSPVITLNSMGRSLTAIOESNOGAFASPIQLORRNGSKFSYDVSTILPIT 252
QY 232 AIMLFVCRDRPSS 244
Db 253 ALMYRCAPPSS 265

RESULT 12
US-08-425-336-1
; Sequence 1, Application US/08425336
; Patent No. 5621083
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroli, Stephen F.
; APPLICANT: Studulka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Rbdsome-Inactivating
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,336
; FILING DATE: 18-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 31394
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-425-336-1

Query Match 26.9%; Score 383; DB 1; Length 267;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;

QY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPL-LRSTIPVSDAQRVIVLITNOGSDSTIAI 61
Db 17 TVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPIN--QRILVELSNHAEISVTLAL 74
QY 62 DVTNLYVAVQAGDOSYFLRDAPDCAE-----RHLPFTGT-TRSSLPFTSGYTDLERYAGH 115
Db 116 -RDQIPLGIEELIQSVSLRY---FGSTRAQARSLIILQIMISANPNPIFMRAROYI 171

Db 75 DVTNAYVGYRAGNSAYFFH--PDNQEDAETHLFTDVQNRRTFAFGNVDRLQLAGN 132
 Qy 116 -RDQIPGIEELIOSVALRY---PGGSTRQAQARSLIILMISSAARFNPITFARROYI 171
 Db 133 LRENIELGNGPLLEAISALYYSTGTQPLTARSFIIICIMISEAARFQYIEGEMRTRI 192
 Qy 172 NSGSEFLPDWYMLELETSMGQOSTOVQSDGVFNNPRLGISTGNFTLSNVRDIASL 231
 Db 193 RYNRSAPDPVSITLSEMSGRISTAIQESNOGAFASPIQLQRRNGSKFSYDVSLIPII 252
 Qy 232 AIMEVCRDRPSS 244
 Db 253 ALMYRCAPPPSS 265

RESULT 13 US-08-488-113B-1

Sequence 1, Application US/08488113B
 Patent No. 5744580
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studulka, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,113B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-488-113B-1

Query Match 26.98: Score 383, DB 1: Length 267;
 Best Local Similarity 39.98: Pred. No. 1e-32;
 Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;
 Qy 4 TGDEYFRPITLLRDYSSGS-FSNEIPL-RQSTIPYSDAQRFLVLTNQGDSITAI 61
 Db 17 TVQSYTFIRAVRKRRLTTGAD/RHEIPLPVRGLPTN--QREIVELSHMAELSVTLAL 74
 Qy 62 DVTNAYVGYRAGNSAYFFH--PDNQEDAETHLFTDVQNRRTFAFGNVDRLQLAGN 132
 Db 75 DVTNAYVGYRAGNSAYFFH--PDNQEDAETHLFTDVQNRRTFAFGNVDRLQLAGN 132
 Qy 116 -RDQIPGIEELIOSVALRY---PGGSTRQAQARSLIILMISSAARFNPITFARROYI 171
 Db 133 LRENIELGNGPLLEAISALYYSTGTQPLTARSFIIICIMISEAARFQYIEGEMRTRI 192
 Qy 172 NSGSEFLPDWYMLELETSMGQOSTOVQSDGVFNNPRLGISTGNFTLSNVRDIASL 231
 Db 193 RYNRSAPDPVSITLSEMSGRISTAIQESNOGAFASPIQLQRRNGSKFSYDVSLIPII 252
 Qy 232 AIMEVCRDRPSS 244
 Db 253 ALMYRCAPPPSS 265

RESULT 14 US-08-477-484B-1

Sequence 1, Application US/08477484B
 Patent No. 5756699
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studulka, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,484B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 03/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 03/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889

Best Local Similarity 35.9%; Pred. No. 6.3e-23;
Matches 85; Conservative 37; Mismatches 107; Indels 8; Gaps 4;

QY 8 YFRFTTLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITAAIDVTNLY 67
Db 17 YVNFLELRLKRLPEGNSHGIPLRKCC--DDPGKCFVLAALNDNGQLAEIAIDVTSY 74
QY 68 VVAYAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126
Db 75 VVGVOYRNRSYFFKADPAAYEGLEFKNTIKTRLHFGSYPSLEGEKAYRETTDLGIEPLR 134
QY 127 --IQSVSALRYPGGSTRAOARSLIILQIMISEARFNPIFWARQYINSGESFLPDMYL 184
Db 135 IGKKLDENAIIDNKPEIASLIVYIOWYSEARFTFENQIRN--NFQQRIRPANNTI 192
QY 185 ELETSMGQOSTQVOOS--TDGVFNPNPRLGISTGNFVTLNVNRDVIASLAIMLFVCHDRPS 243
Db 193 SLENKWKLSFOIRTSANGMFSSEAVLEBRANGKRYVAVDQVKRIALLKFFVDKD 249

RESULT 5

US-09-765-527-259
Sequence 259, Application US/09765527
Patent No. US20020006638A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,527

FILING DATE: 18-Jan-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/621,803

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447

REFERENCE/DOCKET NUMBER: 27129/33199

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

INFORMATION FOR SEQ ID NO: 259:

SEQUENCE CHARACTERISTICS:

LENGTH: 293 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 259:

US-09-765-527-259

Query Match 21.7%; Score 309; DB 10; Length 293;

Best Local Similarity 35.1%; Pred. No. 7.8e-23;

Matches 85; Conservative 40; Mismatches 109; Indels 8; Gaps 4;

QY 8 YFRFTTLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITAAIDVTNLY 67
Db 39 YVNFLELRLKRLPEGNSHGIPLRKCC--DDPGKCFVLAALNDNGQLAEIAIDVTSY 96
QY 68 VVAYAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126

Db 97 VVGVOYRNRSYFFKADPAAYEGLEFKNTIKTRLHFGGTYPGLEGEKAYRETTDLGIEPLR 156

QY 127 --IQSVSALRYPGGSTRAOARSLIILQIMISEARFNPIFWARQYINSGESFLPDMYL 184

Db 157 IGKKLDENAIIDNKPEIASLIVYIOWYSEARFTFENQIRN--NFQQRIRPANNTI 214

QY 185 ELETSMGQOSTQVOOS--TDGVFNPNPRLGISTGNFVTLNVNRDVIASLAIMLFVCHDRPS 243

Db 215 SLENKWKLSFOIRTSANGMFSSEAVLEBRANGKRYVAVDQVKRIALLKFFVDKDRS 274

QY 244 SS 245

Db 275 AA 276

RESULT 6

US-09-765-527-253

Sequence 253, Application US/09765527

Patent No. US20020006638A1

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,527

FILING DATE: 18-Jan-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/621,803

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447

REFERENCE/DOCKET NUMBER: 27129/33199

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

INFORMATION FOR SEQ ID NO: 253:

SEQUENCE CHARACTERISTICS:

LENGTH: 309 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 253:

US-09-765-527-253

Query Match 21.7%; Score 309; DB 10; Length 309;

Best Local Similarity 35.1%; Pred. No. 8.4e-23;

Matches 85; Conservative 40; Mismatches 109; Indels 8; Gaps 4;

QY 8 YFRFTTLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITAAIDVTNLY 67
Db 39 YVNFLELRLKRLPEGNSHGIPLRKCC--DDPGKCFVLAALNDNGQLAEIAIDVTSY 96
QY 68 VVAYAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126
Db 97 VVGVOYRNRSYFFKADPAAYEGLEFKNTIKTRLHFGGTYPGLEGEKAYRETTDLGIEPLR 156
QY 127 --IQSVSALRYPGGSTRAOARSLIILQIMISEARFNPIFWARQYINSGESFLPDMYL 184

```
Db 157 IGTKLDENALIDNKKPELILASLLVLIQVNSEARFTEIENQIRN--NFOQIRIPANNNTI 214
OY 185 ELETSMGQOSTQVOQS--TDCGFNNPRLGISTGNEFTLSNNDVLAISLIMLFVORDRPS 243
Db 215 SLENKMGKLSFQIRTSGANMGSEAVELEBRANKKYYTAVDQVFKIALILKFDKDPKS 274
OY 244 SS 245
Db 275 AA 276

RESULT 7
US-09-765-527-251
; Sequence 251, Application US/09765527
; Patent No. US2002006538A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; Fusion Proteins and BPI-Derived Peptides
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshail, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 251:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-765-527-251
Query Match 21.7%; Score 309; DB 10; Length 332;
Best Local Similarity 35.1%; Pred. No. 9,3e-23;
Matches 85; Conservative 40; Mismatches 109; Indels 8; Gaps 4;

Db 157 IGTKLDENALIDNKKPELILASLLVLIQVNSEARFTEIENQIRN--NFOQIRIPANNNTI 214
OY 185 ELETSMGQOSTQVOQS--TDCGFNNPRLGISTGNEFTLSNNDVLAISLIMLFVORDRPS 243
Db 215 SLENKMGKLSFQIRTSGANMGSEAVELEBRANKKYYTAVDQVFKIALILKFDKDPKS 274
OY 244 SS 245
Db 275 AA 276

RESULT 8
US-09-792-793A-34
; Sequence 34, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggin, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 34
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Bryonia dioica
US-09-792-793A-34
Query Match 20.8%; Score 297; DB 9; Length 247;
Best Local Similarity 33.5%; Pred. No. 9,5e-22;
Matches 80; Conservative 54; Mismatches 91; Indels 14; Gaps 8;

Db 157 IGTKLDENALIDNKKPELILASLLVLIQVNSEARFTEIENQIRN--NFOQIRIPANNNTI 214
OY 185 ELETSMGQOSTQVOQS--TDCGFNNPRLGISTGNEFTLSNNDVLAISLIMLFVORDRPS 243
Db 215 SLENKMGKLSFQIRTSGANMGSEAVELEBRANKKYYTAVDQVFKIALILKFDKDPKS 274
OY 244 SS 245
Db 275 AA 276

RESULT 9
US-09-792-793A-36
; Sequence 36, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggin, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 36
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Momordica charantia
US-09-792-793A-36
Query Match 17.5%; Score 249; DB 9; Length 250;
Best Local Similarity 33.5%; Pred. No. 5,4e-17;
Matches 69; Conservative 28; Mismatches 83; Indels 26; Gaps 5;
```

us-09-627-165e-16.rapp

Page 5

```

1 RESULT 10
2 US-09-792-793A-35
3 ; Sequence 35, Application US/09792793A
4 ; Patent No. US20020168370A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: McDonald, John R.
7 ; APPLICANT: Cogdins, Philip
8 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
9 ; FILE REFERENCE: OTHER INFLAMMATORY CONDITIONS
10 ; CURRENT APPLICATION NUMBER: US/09/792,793A
11 ; CURRENT FILING DATE: 2001-02-22
12 ; NUMBER OF SEQ ID NOS: 93
13 ; SOFTWARE: PatentIn Ver. 2.0
14 ; SEQ ID NO 35
15 ; LENGTH: 275
16 ; TYPE: PRT
17 ; ORGANISM: Saponaria officinalis
18 ; US-09-792-793A-35

```

```

Query Match 13.5%, Score 192, DB 9, Length 275,
Best Local Similarity 26.2%, Pred. No. 2,7e-11,
Matches 75, Conservative 56, Mismatches 99, Indels 56, Gaps 15,

QY      4 TGDVEFRFTLLRDVSSGSPSNEIPLLRQ-----SIFPSDAQRFVLELTNOGDSIT 58
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      12 TAGQSSSFVDKIRNNVKD-----PMLKGTGCDIAVIGPPSKKEFLRINFQSSRG-fvs 63
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      59 AADVTNLYVAAYQAGD-----QSYFLRDAPDGAe-RHfE---TGTRESLEFQSYIDL 109
      | : | | | | | | | | | | : | : | : | : | : | : | : | : | : |
DB      64 LGLKRDNLVYAAYLAMPNTWNNRAYYFRSEITSAESTALFPPEATTANQALAEITYDYOSI 123
      | : | | | | | | | | | | : | : | : | : | : | : | : | : | : |
QY      110 ERYA-----GHRQIDPLGIELLOSVALRIYGGSTRAQASLLITLQMTSEARFNPDI 163
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      124 EKNOITGGDGSRKELGIDLLSTSMWAEKKARVYVCKDEARFLILAIQMTAEARF--- 180
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      164 FWRARQYI-----NSGESFLPDMYLEMLETSMGOOSTQVQ-QSDVGYFNPNFRIGISTG 216
      | : | | | | | | | | | | : | : | : | : | : | : | : | : | : |
DB      181 -----RYIQMLVYNFKNFKFNSENKVIQFPEVMKKSTRAITGDANNGVFNNDDYDFG----- 231
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      217 NEVTLNVNRQVIALMLFVCRDRPS---SSDVR-YWPLVIRPVL 258
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      232 -FGKRVQYKDL--QMGLMYLGRPKSSNEANSTVBHYGPL--KPTL 272
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 11
US-09-792-793A-85
: Sequence 85, Application US/09792793A
: Patent No. US20020168370A1
: GENERAL INFORMATION:
: APPLICANT: McDonald, John R.
: TITLE OF INVENTION: COGGINS, Philip
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
: FILE REFERENCE: 2520-0-601D
: CURRENT APPLICATION NUMBER: US/09/792, 793A

```

```

?
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ. ID NOS: 93
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 85
? LENGTH: 254
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Methionine-Saporin fusion
? US-09-792-793A-85

```

[illegible]

```

RESULT 12
US-09-792-793A-79
; Sequence 79, Application US/09'72793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggin, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 79
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion pr
US-09-792-793A-79

Query Match          12.8%; Score 182; DB 9; Length 327;
Best Local Similarity 26.3%; Pred. No. 3.3e-10;
Matches 68; Conservative 47; Mismatches 96; Indels 48; Gaps 11;

QY      4 TGDEFRITTLRDYVSSGSFSENIPLLRQ-----STIPVSDAORFVLVELTNGGDSIT 58
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      86 TAGOYSSVPDKIRNNVKD-----PNLKYGTDIAIVGPSKKEFLRINQSSRG--TVS 137

QY      59 AADITNLYVAAYAGD----QGYFLRADPDCAE-RHLF---TGTRSSLPTFGSYTDL 109
       | ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     138 LGLRDNLYVAAYLAMONTNNYNRAIYFRSETTSASTALPPEATTANOKALEYTEDYOSt 197

QY    110 ERYA-----GHRDQIDLGIEELIQSVASLARYPGSGTPRAQRSLILIIQWISNAERNPt 163
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

1      RESULT 14
2      US-09-792-793A-73
3      ; Sequence 73, Application US/09792793A
4      ; Patent No. US20020168370A1
5      GENERAL INFORMATION:
6      ; APPLICANT: McDonald, John R.
7      ; APPLICANT: Coggins, Philip
8      ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE ANT
9      ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
10     ; FILE REFERENCE: 25020-601D
11     ; CURRENT APPLICATION NUMBER: US/09/792.793A
12     ; CURRENT FILING DATE: 2001-02-22
13     ; NUMBER OF SEQ ID NOS: 93
14     ; SOFTWARE: PatentIn Ver. 2.0

```

	Query Match	12.8%	Score 182:	DB 9:	Length 332:	
	Best Local Similarity	26.3%	Pred. No.	3.4e-10:		
Matches	68:	Conservative	47:	Mismatches	96:	Gaps 11:
QY	4 TGDEPERFTLLRDVSSGSGFSNEIPLRQ-----STTPVSDAORFYLVELLTNOGDISTT	58	I : :	:	:	: :
Db	91 TAGOISSEFDKTRNNVKD-----PNLKYGCTDIAGPSPKEFLIRINFOSSRG--TVS	142	:	:	:	:
QY	59 AAIDVTNLYVAAYQAAD-----QSIFYLADAADGAE-RHLF---TGTRRSSLPFGGSYTDL,	109	:	:	:	:
Db	143 LGGRKNLVVVYYVLAMDNTNNRRAVFPESETTSASASTALFEAATNAOKALETYEDQSI	202	:	:	:	:
QY	110 ERKA-----GHDDIPLGTIELLIOSVALRYPGGSTRAQRSLILLIOMISEARNPPI	163	:	:	:	:
Db	203 EKNAQTITGGDSKKELGITDIDLSTSMEEVNKKAVRWKVDAREFLIIAIOTMEAARF--	259	:	:	:	:

Qy 164 FWRAROYI-----NSGESEFLPDMYMLEFTSMGOOSTOYQ-QSTDCVENNPRIGISTG 216
Db 260 -----RYQNLVYIKNFENKENSENKYVIOFEVNMKKISTAIYGDANKGVFNKDYDFG----- 310
Qy 217 NEVTLSNVRDYIASLAIML 235
Db 311 -FGKVRQVKDIDQMGILMYL 328

Search completed: March 18, 2003, 08:28:38
Job time : 16.4221 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:55 ; Search time 23.1332 Seconds
(without alignments)
1159.439 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425

Sequence: 1 HQTGDEYFRFTLLRDYVS.....NSGAVDVTCTASEPTVRLV 279

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	78.5	254	2 PD0018	mistletoe lectin I
2	514	36.1	528	1 TZLSA	abrin-a precursor
3	481	33.8	528	2 S32431	abrin-d precursor
4	481	33.8	562	2 S16022	abrin-c precursor
5	471.5	33.1	527	2 S32430	abrin-b precursor
6	458	30.1	251	2 C39761	abrin (clone 7.2)
7	427	29.9	570	2 S62627	agglutinin I precu
8	425.5	29.9	576	1 RLCSAG	ricin D precursor
9	423	29.7	564	1 RLCSAG	agglutinin precus
10	340.5	23.9	286	2 S25560	RNA N-glycosidase
11	336	23.6	289	1 RLUTZ	RNA N-glycosidase
12	333.5	23.4	286	2 JCT425	RNA N-glycosidase
13	331.5	23.3	289	2 JCT506	RNA N-glycosidase
14	329	23.1	247	2 JCT503	karasurin C - Tric
15	327	22.9	247	2 JCT503	karasurin-B - Tric
16	324.5	22.8	245	2 JCT4840	karasurin - Mongol
17	311	21.8	316	2 JCT0753	RNA N-glycosidase
18	296.5	20.8	286	1 RLPU6G	RNA N-glycosidase
19	291.5	20.5	278	2 S23519	beta-luifin - smoo
20	286.5	20.1	277	2 S22494	RNA N-glycosidase
21	280	19.6	250	2 JN0108	luifin-b - smooth
22	249	17.5	278	2 A39817	RNA N-glycosidase
23	203	14.2	313	2 S17757	RNA N-glycosidase
24	202.5	14.2	261	2 JCT401	antifurcal protein
25	186	13.8	294	2 S28421	RNA N-glycosidase
26	187	13.2	253	2 S28539	RNA N-glycosidase
27	187	13.1	283	2 S05205	RNA N-glycosidase
28	186	13.1	253	2 S28542	RNA N-glycosidase
29	184	12.9	272	2 JCT4811	betavulgin - beet

30	182	12.8	253	2 S29331	RNA N-glycosidase
31	182	12.8	292	1 RLQHC2	RNA N-glycosidase
32	180.5	12.7	106	2 B39761	abrin (clone 3.7)
33	179	12.6	253	2 S28541	RNA N-glycosidase
34	161	11.3	253	2 A58923	RNA N-glycosidase
35	157.5	11.1	310	2 S46239	ribosome-inactivat
36	148.5	10.4	293	2 S17519	RNA N-glycosidase
37	143.5	10.1	236	2 S17932	RNA N-glycosidase
38	134	9.4	289	2 T12573	RNA N-glycosidase
39	129	9.1	319	2 S58343	Shiga-like toxin I
40	128.5	9.0	319	2 S01032	Shiga-like toxin I
41	128	9.0	319	2 S21940	hypothetical prote
42	128	9.0	319	2 E90779	Shiga toxin 2 subu
43	128	9.0	319	2 G85640	Shiga toxin 2 subu
44	127.5	8.9	320	2 S42608	Shiga-like toxin -
45	127	8.9	319	2 I76713	variant shiga-like

ALIGNMENTS

RESULT 1

PD0018 mistletoe lectin I A chain - Viscum album (fragment)

C:Species: Viscum album

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 26-Aug-1999

C:Accession: PD0018

R:Eschenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, A.

Biochem. Biophys. Res. Commun. 247, 367-372, 1998

A:Title: Primary structure and molecular modeling of mistletoe lectin I from Viscum A

A:Reference number: PD0018; MUID:98308123; PMID:9642133

A:Accession: PD0018

A:Molecule type: protein

A:Residues: 1-254 <ESC>

C:Superfamily: ricin; RNA N-glycosidase homology

F:7-246/Domain: RNA N-glycosidase homology <RNG>

Query Match 78.5%; Score 1119; DB 2; Length 254;

Best Local Similarity 89.0%; Pred. No. 1.6e-93;

Matches 218; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY	1	HQTGDEYFRFTLLRDYVS	SFSFSENIPLRQSTIPVSDAQRFLVLTNGQSDSITAA	60
DB	10	HQTGDEYFRFTLLRDYVS	SFSFSENIPLRQSTIPVSDAQRFLVLTNGQSDSITAA	69
QY	61	IDVTMLVYVAYAGQSTFL	RDAPGARHLFTGTRSSLPFTGSYTDLERYAGHRQIP	120
DB	70	IDVTNAYVAYAGQSTFL	RDAPGARHFTGTRSSLPFTGSGYTDLERYAGHRQIP	129
QY	121	IGIEELIOSVSLARYPGS	TRQAARSLITLIOMISEARFNPIMFARQYINSGESFLPD	180
DB	130	IGIDIDLIOSVSLARYP	GSSTRQAARSLITLIOMISEARFNPIMFARQYINSGESFLPD	189
QY	181	MYMLELTSWGOOSTQVO	STDGVEFNNPFRLGISTGNFVLSNRYDIASLAIMFLPCRD	240
DB	190	YVMELELTSWGOOSTQ	VGHSITDGVENNPRIALIPGNGVVTITNNRDVIAISLAIMFLPCGE	249
QY	241	RRSSS	245	
DB	250	RRSSS	254	

RESULT 2

TZLSA

abrin-a precursor - Indian licorice (fragment)

N:Contains: RNA N-glycosidase (EC 3.2.2.22)

C:Species: Abrus precatorius (Indian licorice)

C:Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999

C:Accession: S32429; JTO202; A39761; JCT1398; S14472; S24133; S74110; S74111

R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.

J. Mol. Biol. 229, 263-267, 1993

A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.

A:Reference number: S32429; MUID:913132798; PMID:8421313

A:Accession: S32429
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'E', 2-528 <HUN>
A:Cross-references: GB:M8344; NID:g166294; PIDN:AAA3624.1; PID:g166295
A:Note: The coding region for the sequence shown is preceded by an ATG codon
A:Note: residues 1-8 were derived from the synthesized primer
R:Funatsu, G.; Taguchi, Y.; Kamenosono, M.; Yanaka, M.
AGRIC. Biol. Chem. 52, 1095-1097, 1988
A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from
A:Reference number: J70202
A:Accession: J70202
A:Molecule type: protein
A:Residues: 1-201,203-251 <FUN>
A:Note: The amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329; PMID:2016300
A:Accession: A39761
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 'E', 2-251 <EVE>
A:Cross-references: GB:X54872
A:Note: residues 1-8 were derived from the synthesized primer
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
A:Reference number: J01398; MUID:93169023; PMID:7763422
A:Contents: seeds
A:Accession: J01398
A:Molecule type: protein
A:Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>
A:Experimental source: seed
R:Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'ME', 2-251 <EV2>
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: EMBL:X54873; NID:g16090; PIDN:CA38655.1; PID:g16091
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.
PDBS Lett. 309, 115-118, 1992
A:Title: The complete primary structure of abrin-a B chain.
A:Reference number: S24133; MUID:92371656; PMID:1505674
A:Accession: S24133
A:Molecule type: protein
A:Residues: 262-297, 'Y', 999-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
Eur. J. Biochem. 240, 564-569, 1996
A:Title: Probing the domain structure of abrin-a by tryptic digestion.
A:Reference number: S74110; MUID:97008945; PMID:8856055
A:Accession: S74110
A:Molecule type: protein
A:Residues: 89-108,154-172 <LIN>
A:Experimental source: seed
A:Accession: S74111
A:Molecule type: protein
A:Residues: 262-276, 'X', 278-280, 329-348, 369-388, 399-418 <LIM>
A:Experimental source: seed
C:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inhibits
protein synthesis by inhibiting the eukaryotic 80S ribosome, and a B chain, which is
responsible for binding to the cell surface. The A and B chains are linked by a single disulfide
bond.
C:Superfamily: ricin; RNA N-glycosidase homology
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglyltamic acid;
F:1-251/Product: abrin-a chain A #status experimental <ACH>
F:7-246/Domain: RNA N-glycosidase homology <RNG>
F:261-528/Product: abrin-a chain B #status experimental <BOC>
F:283-325,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Glu) #status experimental
F:774,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted

F;247-269,286-305,329-346,417-430,456-473/disulfide bonds: #status predicted
F;288,312/binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F;361,401/binding site: carboxylate (Asn) (covalent) #status experimental
F;500,521/binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 36.1%; Score 514; DB 1; Length 528;
Best Local Similarity 42.5%; Pred. No. 2,1e-38;
Matches 122; Conservative 45; Mismatches 88; Indels 32; Gaps 8;

QY 4 TGEDEFREITLLRDVSSGFSNFNEIPLRKOSTPIPVSDAQRFVELTNGOGDSITAAIDV 63
 :|||:::~::~:|||:::~::~|||||:::~::~|||
Db 13 TSQS>KQGTEALRERL-RGLIHDPVLDPDT-VLEGRRIYTEVLESNDTESIEVGDI 70

QY 64 TNLYAVAYQAQDQSYFLFDAPDGAEHLFTTTSSSLPFTGSYTDLERYAGH-RDQIPLG 122
 |||||:::~::~|||||:::~::~|||||:::~::~|||||
Db 71 TNAYVAARAQTOSFYFLDAPSASADYLFTGTDDHSLPFEVTYDDELMWAQSQIPLG 130

QY 123 IEELIOSVALRYFGSGSTRACARSLILTIOMISEAFNPFPMRAFORYINSEFSLPDMY 182
 ::|||:::~::~|||||:::~::~|||||:::~::~|||||
Db 131 LQALTHGISFFRRSGGNDEEAKRLIVITIOWALEARRTISNRVRSIGTGTFQPDPA 190

QY 183 MLEETSWGQOSTOVQOYSTDVGFNNFPRLISTGNFNLSNVWD-----VIASL 231
 :|||:::~::~|||||:::~::~|||||:::~::~|||||
Db 191 MISLENMNDNLIRGVQESVDPTFPNQ-----VLTFRIRNEPVIDLSHPYA VL 240

QY 232 AIMLFVCNDRPSSSDVRKYWPLYIRPVLENSCAVDVYTCTASEPYVRI 278
 :|||:::~::~|||||:::~::~|||||:::~::~|||||
Db 241 ALMFVFC-NPPNAMOS---PLLIRSIVEKS---KICSSREPEYVRI 279

RESULT 3
S32431
abrin-d precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
C:Accession: S32431; S34408
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A>Title: Primary structure of three distinct isoabrins determined by cDNA sequencing .
A:Reference number: S32429; MUID:93133798; PMID:8421313
A:Accession: S32431
A:Molecule type: mRNA
A:Residues: 1-528 <HUN>
A:CROSS-references: GB:M98346
R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
submitted to the EMBL Data Library, March 1993
A:Reference number: S34408
A:Accession: S34408
A:Molecule type: mRNA
A:Residues: 1-169,'C',171-320,'L',322-528 <HU2>
A:CROSS-references: GB:M98346
C:Comment: Abirin consists of an A chain, which inhibits protein synthesis by inactivating
the A and B chains are linked by a single disulfide bond, which is essential for tox
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin
F;1-251/Product: abrin-d chain A #status predicted <ACH>
F;7-246/Domains: rRNA N-glycosidase homology <RNC>
F;261-528/Product: abrin-d chain B #status predicted <BCH>
F;283-335,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F;74,113,195,196/binding site: substrate (TYR, TYR, Glu, Asn) #status predicted
F;164,167/active site: Glu, Arg #status predicted
F;200,253,361,401,402/binding site: carbonylate (Asn) (covalent) #status predicted
F;247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted
F;288,312/binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F;500,521/binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 33.8%; Score 481; DB 2; Length 528;
Best Local Similarity 43.7%; Pred. No. 2,1e-35;
Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

4 TGDEFREITTLRDVSSGFSNFNEIPLRKOSTPIPVSDAQRFVELTNGOGDSITAAIDV 63


```
Db 13 TSQSYKQFEALRQRL-TGGLIHDPVLPDPT-TVEERNRIITVELNSERESELEVIDV 70
QY 64 TNLVYVAOAGDSYFLRDPADGAEHLFTGTRSSLPFTGSDYDLERYAGH-RDQIPLG 122
Db 71 TNAVYVAIRAGOSYFLRDPASASTYFLPCTQRYSLRFDOSYDDLBRMAHQTRREISLG 130
QY 123 IEELIOSVSALRYPGSGSTRQAARSLITLIOMISEAARFNPFWMARQYINSGESFLPDMY 182
Db 131 LQALTHAISFLRSGASNDKAKRTLIYIOMASEAARFYISNRGVSIKRTGTAFOQDPA 190
QY 183 MLELETSWGQOSTVOQSTDGVE-NNPFRLGISTGNFVTLISNVDYIASLAIMLFCRDR 241
Db 191 MLELNMMNDNLSSGVQOSYQDPTFPNNVILSSINQPVVDSLSHPYAVIALMLFVC-NP 249
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
Db 250 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 279
```

RESULT 4

```
S16022
abrin-c precursor - Indian licorice
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S16022
R:Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Platak, M.
Eur. J. Biochem. 198, 723-732, 1991
A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-chain
A:Reference number: S16022; MUID:91266957; PMID:2050349
A:Accession: S16022
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <MOO>
A:Cross-references: EMBL:X55667; NID:g16084; PIDN:CAA39202.1; PID:g16085
C:Comment: Abirin consists of an A chain, which inhibits protein synthesis by inactivating
the A and B chains are linked by a single disulfide bond, which is essential for toxicity
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid
F:35-285/Product: abrin-c chain A #status predicted <ACH>
F:41-280/Domain: rRNA N-glycosidase homology <RNG>
F:295-562/Product: abrin-c chain B #status predicted <BCH>
F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats
F:53/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:198,201/Active site: Glu, Arg #status predicted
F:234,287,395,435,436/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:281,303,320-338,383-380,451-464,490-507/Disulfide bonds: #status predicted
F:322,346/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
F:534,555/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
```

Query Match 33.8% Score 481; DB 2; Length 562;

Best Local Similarity 43.7%; Pred. No. 2,2e-35;

Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

```
QY 4 TGEYFRFTLLRDYVSSGFSNEIPLRQSTIPVSDAQRFLVELTNGGDSITAAIDV 63
Db 47 TSQSYKQFEALRQRL-TGGLIHDPVLPDPT-TVEERNRIITVELNSERESELEVIDV 104
QY 64 TNLVYVAOAGDSYFLRDPADGAEHLFTGTRSSLPFTGSDYDLERYAGH-RDQIPLG 122
Db 105 TNAVYVAIRAGOSYFLRDPASASTYFLPCTQRYSLRFDOSYDDLBRMAHQTRREISLG 164
QY 123 IEELIOSVSALRYPGSGSTRQAARSLITLIOMISEAARFNPFWMARQYINSGESFLPDMY 182
Db 165 LQALTHAISFLRSGASNDKAKRTLIYIOMASEAARFYISNRGVSIKRTGTAFOQDPA 224
QY 183 MLELETSWGQOSTVOQSTDGVE-NNPFRLGISTGNFVTLISNVDYIASLAIMLFCRDR 241
Db 225 MLELNMMNDNLSSGVQOSYQDPTFPNNVILSSINQPVVDSLSHPYAVIALMLFVC-NP 283
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
```

```
Db 284 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 313
```

RESULT 5

```
S32430
abrin-b precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
C:Accession: S32430; JCI399
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32430
A:Molecule type: mRNA
A:Residues: 1-527 <HUN>
```

```
A:Cross-references: GB:M98345; NID:g166296; PIDN:AAA32625.1; PID:g166297
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, to
A:Reference number: JCI398; MUID:93169023; PMID:7763422
A:Accession: JCI399
A:Molecule type: protein
A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-
A:Experimental notes: seed
```

```
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin
F:1-250/Product: abrin-b chain A #status predicted <ACH>
F:7-245/Domain: rRNA N-glycosidase homology <RNG>
F:260-527/Product: abrin-b chain B #status experimental <BCH>
F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F:74,113,119,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:110,360,400/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:163,166/Active site: Glu, Arg #status predicted
F:267,268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
F:287,311/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
F:499,520/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
```

Query Match 33.1% Score 471.5; DB 2; Length 527;

Best Local Similarity 42.6%; Pred. No. 1.5e-34;

Matches 118; Conservative 46; Mismatches 100; Indels 13; Gaps 8;

```
QY 4 TGEYFRFTLLRDYVSSGFSNEIPLRQSTIPVSDAQRFLVELTNGGDSITAAIDV 63
Db 13 TSQSYKQFEALRQRL-TGGLIHDPVLPDPT-TQERNRIITVELNSDSTESIEAGIDV 70
QY 64 TNLVYVAOAGDSYFLRDPADGAEHLFTGTRSSLPFTGSDYDLERYAGH-RDQIPLG 122
Db 71 TNAVYVAIRAGRSYFLNDAPTSASRYLFTGQOYSLRFSNSYIDLLEIRLAQTRQOQIPLG 130
QY 123 IEELIOSVSALRYPGSGSTRQAARSLITLIOMISEAARFNPFWMARQYINSGESFLPDMY 182
Db 131 LQALRHAISFLQ-SGTDQOELARTLIYIOMASEAARFYISNRGVSIKRTGTAFOQDPA 189
QY 183 MLELETSWGQOSTVOQSTDGVE-NNPFRL-GISTGNFVTLISNVDYIASLAIMLFCRDR 241
Db 190 MLELNMMNDNLSSGVQOSYQDPTFPNAVTLRSVNNQPVIVDSLIHQSVAVIALMLFVC-NP 248
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
Db 249 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 278
```

RESULT 6

```
C39761
abrin (clone 7.2) precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C:Accession: C39761; S14471
```

	Query Match	30.0%	Score 427;	DB 2;	Length 570;
	Best Local Similarity	37.6%	Pred. No. 1,7e-30;		
	Matches 109; conservative	51;	Mismatches 110;	Indels 20;	Gaps 9
OY	3 TTGDEYRITLLRDVYSGFSN-EIPLRSTIPYSQAQRVLVELNOGSDTIAI	61			
	: : : : :				
Db	42 TCADYYEPLRLAOEVILGNHTADLPVLNES-OVSNSNRKVALPILNDSDDVTAI	100			

A:Accession: A03374
A:Molecule type: Protein
A:Residues: 315-333, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 'F', '<UN>
507, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F', '<UN>
A:Note: This paper, one of a series, summarizes the experimental details for the detaile
R:Ready, W.P.; Kim, Y.; Robertus, J.D.

R: Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, L.
submitted to the Brookhaven Protein Data Bank, July 1994

activation activities. It is capable of sensing...

F;4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 22.9%; Score 327; DB 2; Length 247;

Best Local Similarity 37.2%; Pred. No. 5.8e-22;

Matches 89; Conservative 51; Mismatches 85; Indels 14; Gaps 9;

```
OY 4 TGDEYFRITLLRDYVSSGFSFSEIPLLRQSTIPVSDAQRFEVLVELTNGGDSITRAIDV 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 TSSYGVFISNLKRALPYERKLYDIPLLR-STLPGS--GRYALHLITNYADETISVAIDV 66
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 64 TNLVYVAVQAGDOSYFLDA-PDGAERHLFTGTRS-SLPETGSYTDLERVAGH-RDQIP 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 TNYVYVAGYFAGDTSYFNEASATEAKYVEKDAKRYKVTLPYSGNYERLQIAAGKIRENIP 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 LGIEELIGSVSALRRYPGSGSTRQAASLIIITOMISEARFNPFPWRARQYINSGESFLPD 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 LGDPLDLSAITTLFFYNNNSAASA-LMWLIQSTSEARARYKFTEDQIGKRVN--KTFLPS 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 181 MYMLELETSMGOOSTQVQ--OSTDGVENNPFRLGISTGNFVTLNSVRD--VIASLAIML 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 183 LATISLENSWSALSQIQIASTNNGQFEPVVLINQONQRTITNVDAQVVTSMIALLL 241
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: March 18, 2003, 08:27:32
Job time : 25.1332 secs

GenCore version 5.1.4_p5_A578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:01:23 ; Search time 12.6181 seconds
(without alignments)
917.088 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1423
Sequence: 1 HQTGDEYFRFTILLRDYVS.....NSGAVDVTCYASEPIRVIV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	78.5	254	1 MIA_VISAL	P81446 Viscum albu
2	514	36.1	528	1 ABRA_ABRPR	P11140 abrus preca
3	481	33.8	562	1 ABRA_ABRPR	P28590 abrus preca
4	471.5	33.1	527	1 ABRA_ABRPR	P06077 abrus preca
5	425.5	29.9	576	1 RIC1_RICCO	P02879 ricinus com
6	423	29.7	564	1 AGGL_RICCO	P06750 ricinus com
7	393	27.6	563	1 NIGB_SAMNT	P35183 sambucus ni
8	354	24.8	282	1 RIP2_BRIDI	P33183 bromelia dio
9	340.5	23.9	286	1 RIP2_MOMBA	P29339 momordia b
10	336	23.6	289	1 RIP1_TRIKI	P09989 trichosanthe
11	331.5	23.3	289	1 RIP1_TRIKI	P24478 trichosanthe
12	312.5	21.9	294	1 RIP1_TRIKI	P56626 trichosanthe
13	311	21.8	316	1 RIPG_GELMU	P33186 gelonium mu
14	308.5	21.6	286	1 RIP1_GUCFI	O96fx4 cucumis fig
15	297.5	20.9	290	1 RIP1_BRIDI	P33183 bromelia dio
16	296.5	20.8	286	1 RIP1_MOMCH	P16094 momordia c
17	286.5	20.1	277	1 RIP1_LURCY	P22851 luffa cylin
18	280	19.6	250	1 RIPB_LURCY	P22851 luffa cylin
19	249	17.5	278	1 RIPB_MIRJA	P21326 mirabilis j
20	203	14.2	313	1 RIP1_PHYAM	P10297 phytoacca
21	202.5	14.2	261	1 RIP1_PHYAM	P23339 phytoacca
22	196	13.8	294	1 RIP6_PHYAM	O03464 phytoacca
23	192	13.5	299	1 RIP6_SAPOF	P20656 saponaria o
24	188	13.2	253	1 RIP5_SAPOF	O41389 saponaria o
25	186	13.1	253	1 RIP7_SAPOF	O41391 saponaria o
26	182	12.8	292	1 RIP2_SAPOF	P27559 saponaria o
27	157.5	11.1	310	1 RIP2_PHYAM	O40772 phytoacca
28	148.5	10.4	293	1 RIP3_DIACA	P24476 dianthus ca
29	143.5	10.1	236	1 RIP3_SAPOF	P27560 saponaria o
30	128	9.0	319	1 STRA_BP33	P09388 bacterioph
31	120.5	8.5	315	1 STRA_BP19	P08028 bacterioph
32	120.5	8.5	315	1 STRA_BP30	P10149 bacterioph
33	120	8.4	280	1 RIP1_HORVU	P22244 hordeum vul

34	113	7.9	280	1 RIP2_HORVU	P04399 hordeum vul
35	103.5	7.3	560	1 J160_HORVU	O00531 hordeum vul
36	92.5	6.5	468	1 PPAD_YEAST	P52290 saccharomyc
37	92	6.5	574	1 YYP3_CABEL	P52716 caenorhabdl
38	90.5	6.4	400	1 B3AR_MOUSE	P25962 mus musculu
39	89.5	6.3	351	1 B3AR_CAVPO	O60483 cavla porce
40	88.5	6.2	971	1 SECS_YEAST	P69102 saccharomyc
41	88	6.2	467	1 B1AR_SHEEP	O28927 ovis aries
42	87.5	6.1	400	1 B3AR_RAT	P26256 rattus norv
43	86	6.0	600	1 MDCD_SALTY	P33902 salmonella
44	85.5	6.0	405	1 B3AR_CANFA	O02662 canis famli
45	85	6.0	454	1 DAT_HAEIN	P44951 h diaminobu

ALIGNMENTS

RESULT 1	ID	MLA_VISAL	STANDARD:	PRT:	254 AA.
AC	P81446;	15-DEC-1998 (rel. 37, Created)			
DT	15-DEC-1998 (rel. 37, Last sequence update)				
DT	15-JUN-2002 (rel. 41, Last annotation update)				
DE	Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).				
OS	Viscum album (European mistletoe).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
OC	Santalales; Viscaceae; Viscum.				
OX	NCBI_TaxID=3972;				
RN	[1]				
RP	SEQUENCE.				
RC	STRAIN-Subsp. album;				
RX	MEDLINE=97134581; PubMed=6960141;				
RA	Huguet, Soler M., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,				
RA	Voelter W.;				
RT	"Complete amino acid sequence of the A chain of mistletoe lectin I.";				
RL	FEBS Lett. 399:153-157(1996).				
CC	-1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGLUTINATION (LECTIN ACTIVITY).				
CC	-1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.				
CC	-1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.				
CC	-1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA AND NON-GLYCOSYLATED FORM MLA'.				
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.				
DR	HSPSP; P11140; IABR.				
DR	Interpro: IPR001574; RIP.				
DR	Pfam: PF00161; RIP.				
DR	PRINTS: PR00136; SHTGARICIN.				
DR	PROSITE, PS00275: SHTGA, RICIN; FALSE_NEG.				
KW	Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.				
FT	ACT_SITE 165 .. 165				
FT	CANBOHYD 112 .. 112				
FT	VARIANT 15 .. 15				
FT	VARIANT 15				
FT	VARIANT 112				
FT	VARIANT 66				
FT	VARIANT 112				
FT	VARIANT 116				
FT	VARIANT 133				
FT	VARIANT 134				
FT	VARIANT 140				
FT	VARIANT 144				
FT	VARIANT 151				
FT	VARIANT 179				
FT	VARIANT 184				
FT	VARIANT 184				
FT	VARIANT 190				
FT	VARIANT 218				
FT	VARIANT 223				
FT	VARIANT 224				

FT	VARIANT	231	231	T -> S (IN MLA').
FT	SEQUENCE	235	285	D -> S (IN MLA').
SQ	REFERENCE	9BD3 AA:	8A478 MM;	53BAF98D3ED0FFEE67 CRC64;
	Query Match		78.5%;	Score 1119;
	Best Local Similarity		89.0%;	Pred. No. 3,4e-93;
	Matches 218;	Conservative	11;	Mismatches 16;
				Indels 0;
				Gaps 0;
QY	1	HQTGTGDEYERFTLLRDYVSSGSFSFNEPILLRKOSTIPYSDAQRFVLVELTNOGSDSITAA	60	
		: : : :		
Dd	10	HQTGTGEYERFTLLRDYVSSGSFSFNEPILLRKOSTIPYSDAQRFVLVELTNOGDSVTAA	69	
QY	61	IDWTNLYVVVYOAGDOSYFLRDAPOGAERHFLTGTTRSSLPTGSYTDLEKYAAHRQIP	120	
		: : : :		
Dd	70	IDWTNAVVVYOAGDOSYFLRDAPRGAERHFLTGTTRSSLPTGSGSYPLERYAACHRDQIP	129	
QY	121	LGIIEELIOSVALRYPGSGSTRAOANSLILLIOMISMAAFNPILFRARROYNSNGESFLPD	180	
		: : : :		
Dd	130	LGIDOLLIOSVALNFPGSGSTRTOARSILLIILLOMISEPAEFNLIMRKROYLNSGASFLPD	189	
QY	181	MYMELETSMWCOOSTQYQOOSTDGVFNPNRPGLTGSTENFTLSNVADVIALSLAIMLVCRD	240	
		: : : : :		
Dd	190	VMYLELETSMWCOOSTQOVGHSDGVFNPIRLAIPLRGNYFTLTNVNRDYIASLAIMLFVCGE	249	
QY	241	RPSSS 245		
Dd	250	RPSSS 254		

RESULT 2	
ABRA_ABRPR	
ID ABRA_ABRPR	STANDARD;
	PRT;
	528 AA

AC 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)
 (EC 3.2.2.22); Abrin-a B chain]
 OS *Abrus precatorius* (Indian licorice) (Crab's eye).
 OC Spermatophyta: Viridiplantae: Streptophyta; Embryophyta: Tracheophyta;
 OC Eukaryota: Magnoliophyta: eudicotyledons: core eudicots: Rosidae.
 OC eurosids I: Fabales: Fabaceae: Papilionoideae; Abreae; Abrus.
 NCBI_Taxid=3816;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93132798; PubMed=8421313;
 RX Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "primary structure of three distinct isoabtrins determined by cDNA
 sequencing. Conservation and significance";
 RL J. Mol. Biol. 225:263-267(1993).
 RN (2)
 RP SEQUENCE OF 1-251.
 RA TISSUE=Seed;
 RC Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
 RX "the complete amino acid sequence of the A-chain of abrin-a, a toxic
 RT protein from the seeds of *Abrus precatorius*.";
 RL Agric. Biol. Chem. 52:1095-1097(1988).
 RN (3)
 RP SEQUENCE OF 1-251 FROM N.A.
 RA TISSUE=Leaf;
 RC MEDLINE=91201329; PubMed=2016300;
 RX Evensen G., Mathiesen A., Sundan A.;
 RA "Direct molecular cloning and expression of two distinct abrin
 RT A-chains.";
 RL J. Biol. Chem. 266:6848-6852(1991).
 RN (4)
 RP SEQUENCE OF 262-528.
 RA MEDLINE=92371656; PubMed=1505674;
 RX Chen Y.-L., Chow L.-P., Tsungta A., Lin J.-Y.;
 RA "the complete primary structure of abrin-a B chain.";
 RL FEBS Lett. 309:115-118(1992).
 RN (5)
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).

RX	MEDLINE=9533318; PubMed=7608980;
RA	Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
RL	*Crystal structure of abrin-a at 2.14 A.";
J.	J. Mol. Biol. 250:354-367(1995).
-I-	FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC	ABRIN-A IS MORE TOXIC THAN RICIN.
CC	-I- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
CC	-I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
CC	-I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC	-I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC	-I- INACTIVATING: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC	-I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb.sib.ch).
CC	or send an email to license@isb.sib.ch).
DR	EMBL; M98344; AAA32624.1; ALT_INT.
DR	EMBL; X54872; NOT_ANNOTATED_CDS.
DR	PIR; S34259; TZLSA.
DR	PIR; S24133; S24133.
DR	PDB; IABR; 07-FEB-95.
DR	InterPro: IPRO01574; RIP.
DR	InterPro: IPRO00772; Ricin_B_lectin.
DR	Pfam; PR00161; RIP; 1.
DR	Pfam; PF00652; Ricin_B_lectin; 6.
DR	PRINTS; PR00396; SHIGARICIN.
DR	SMART; SM00458; RICIN; 2.
DR	PROSITE; PS50231; RICIN_B_LLECTIN; 2.
DR	PROSITE; PS00275; SHIGA_RICIN; 1.
KW	Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; Glycoprotein; Lectin; 3D-structure.
KV	
FT	CHAIN 1 251
FT	PEPTIDE 252 261 LINKER PEPTIDE.
FT	CHAIN 262 528 ABRIN-A B CHAIN.
FT	DOMAIN 273 400 RICIN B-TYPE LECTIN 1.
FT	DOMAIN 403 527 RICIN B-TYPE LECTIN 2.
FT	REPEAT 283 325 1-ALPHA.
FT	REPEAT 326 366 1-BETA.
FT	REPEAT 369 401 1-GAMMA.
FT	REPEAT 414 449 2-ALPHA.
FT	REPEAT 453 492 2-BETA.
FT	REPEAT 495 528 2-GAMMA.
FT	ACCL_SITE 164 164 BY SIMILARITY.
FT	DISULFID 247 269 INTERCHAIN (BY SIMILARITY).
FT	DISULFID 286 305 BY SIMILARITY.
FT	DISULFID 329 346 BY SIMILARITY.
FT	DISULFID 417 430 BY SIMILARITY.
FT	DISULFID 456 473 BY SIMILARITY.
FT	MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT	CARBHYD 361 361 N-LINKED (GLCNAC. . .)
FT	CARBHYD 401 401 N-LINKED (GLCNAC. . .)
FT	CONFLICT 202 202 MISSING (IN REF. 2).
FT	CONFLICT 298 298 N->Y (IN REF. 4).
FT	CONFLICT 427 427 M->L (IN REF. 4).
FT	CONFLICT 467 467 T->P (IN REF. 4).
FT	CONFLICT 483 483 V->I (IN REF. 4).
SO	SEQUENCE 528 AA; 59243 MW; ALF7GBECD59BA827 CRC64;
Query Match	Score 514; DB 1; Length 528;
Best Local Similarity	42.5%; Pred. NO. 1, Be-36;
Matches 122; Conservative	45; Indels 32; Gaps 8;

RX MEDLINE-93132798; PubMed-8421313;
 RA Hung C.-R., Lee M.-C., Lee J.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoforms determined by cDNA
 RT sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267 (1993).
 RN [2]
 RP SEQUENCE OF 260-527.
 RC TISSUE-Seed.
 RA MEDLINE-93169023; PubMed-7763422;
 RA Kimura M., Sumizawa T., Funatsu G.;
 RT "The complete amino acid sequences of the B-chains of abrin-a and
 RT abrin-b, toxic proteins from the seeds of *Abrus precatorius*.";
 RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; M98345; AAA32625.1; .
 DR HSSP; P11140; IABR.
 DR InterPro; IPR001574; RIP.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR Pfam; PF00161; RIP; 1.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR PRINTS; PR00396; SHICARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 DR Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin.
 KW CHAIN 1 250
 FT PEPTIDE 251 260 ABRIN-B A CHAIN.
 FT CHAIN 261 260 LINKER PEPTIDE.
 FT DOMAIN 272 399 ABRIN-B B CHAIN.
 FT DOMAIN 402 526 RICIN B-TYPE LECTIN 1.
 FT REPEAT 282 324 RICIN B-TYPE LECTIN 2.
 FT REPEAT 325 365 1-ALPHA.
 FT REPEAT 368 400 1-BETA.
 FT REPEAT 413 448 1-GAMMA.
 FT REPEAT 452 481 2-ALPHA.
 FT REPEAT 494 527 2-BETA.
 FT ACT_SITE 163 163 2-GAMMA.
 FT DISULFID 246 268 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 285 304 BY SIMILARITY.
 FT DISULFID 328 345 BY SIMILARITY.
 FT DISULFID 416 429 BY SIMILARITY.
 FT DISULFID 455 472 BY SIMILARITY.
 FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY
 FT SIMILARITY).
 FT CARBOXYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 282 282 N -> D (IN REF. 2).
 FT CONFLICT 291 291 D -> N (IN REF. 2).

FT CONFLICT 350 351 AE -> PQ (IN REF. 2).
 FT CONFLICT 378 378 S -> N (IN REF. 2).
 FT CONFLICT 426 426 Y -> M (IN REF. 2).
 FT CONFLICT 428 428 L -> D (IN REF. 2).
 FT CONFLICT 431 431 N -> S (IN REF. 2).
 FT CONFLICT 484 484 R -> K (IN REF. 2).
 FT CONFLICT 491 491 N -> S (IN REF. 2).
 FT CONFLICT 493 493 H -> Y (IN REF. 2).
 FT CONFLICT 502 502 R -> Q (IN REF. 2).
 FT CONFLICT 509 509 E -> Q (IN REF. 2).
 FT CONFLICT 513 513 H -> W (IN REF. 2).
 FT CONFLICT 516 516 H -> T (IN REF. 2).
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;
 Query Match 33.1%; Score 471.5; DB 1; Length 527;
 Best Local Similarity 42.6%; Pred. No. 1,2e-34;
 Matches 118; Conservative 46; Mismatches 100; Indels 13; Gaps 8;
 QY 4 TGDEYFETLLRDYVSSGSGSFNEPILRSTIPVSPAOEFVVELTNGCGSTPAIDV 63
 DB 13 TSQYKOFITALRQRL-TQGLHIGIPVLPDPT-TLQERNYIVELSNEDTSTIAGIDV 70
 QY 64 TNLVVAVQAGDQSYFLRDAPDAERHLFTGTRSSLPFTGSTDLERYAGH-RQQIPLG 122
 DB 71 SNAVVAVRAGNSRYFLRDAPDAERHLFTGTRSSLPFTGSTDLERYAGH-RQQIPLG 130
 QY 123 IEEILQSVSLARPGSGSTPAARSLIILQIMISEAFNFIPFARAOYINSGESFLPDMY 182
 DB 121 LQALRAHISFLQ-SGTDDQELARLILVLIQMAEPARYRISYVAGSIRYWTAFQPDAA 189
 QY 183 MLEETSGQSGOQVQOQSGDVFENPRL-GISGNGVLTLSNVDYASIALMFCVDR 241
 DB 190 MISLENMMDLGSQVQOQSGDVFENPRL-GISGNGVLTLSNVDYASIALMFCVDR 248
 QY 242 PSSSDVIRPVLVIRPVLENAGVDDVYTTASEPVRI 278
 DB 249 PNAQOS---PLILRSTVERKS---KICSSRYEPVRI 278
 RESULT 5
 RICIN_RICCO STANDARD: PRT; 576 AA.
 AC P02879; P02880;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ricin precursor [contains: Ricin A chain (rRNA N-glycosidase)
 DE (EC 3.2.2.22); Ricin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
 NCBI_TaxID=3988;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86067214; PubMed-2999712;
 RA Halling K.C., Halling A.C., Murray E.F., Ladin B.F., Houston L.L.,
 RA Weaver R.F.;
 RT "Genomic cloning and characterization of a ricin gene from *Ricinus*
 RT communis.";
 RL Nucleic Acids Res. 13:8019-8033(1985).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92163016; PubMed-1371405;
 RA Tregear J.W., Roberts L.M.;
 RT "The lectin gene family of *Ricinus communis*: cloning of a functional
 RT ricin gene and three lectin pseudogenes.";
 RL Plant Mol. Biol. 18:515-525(1992).
 [3]
 RP SEQUENCE OF 12-576 FROM N.A.
 RX MEDLINE-85179479; PubMed-3838723;
 RA Lamb A., Roberts L.M., Lord J.M.;
 RT "Nucleotide sequence of cloned cDNA coding for preproricin.";

RL Eur. J. Biochem. 148:265-270(1985).
RN
R14] SEQUENCE OF 36-302.
RA Yoshitake S., Funatsu G., Funatsu M.;
R7 "Isolation and sequences of peptic peptides, and the complete
R7 sequence of the chain of ricin-D.";
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN
R5] SEQUENCE OF 315-576.
RA Funatsu G., Kimura M., Funatsu M.;
R7 "Primary structure of the chain of ricin D.";
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN
R6] CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RA MEDLINE-90344223; PubMed-3168517;
R7 Kimura Y., Kusuhara H., Tada M., Takagi S., Funatsu G.;
R7 "Structural analyses of sugar chains from ricin A-chain variant.";
RL Agric. Biol. Chem. 54:157-162(1990).
RN
R7] REVIEW.
RA MEDLINE-21480122; PubMed-11595634;
R7 Olsnes S., Kozlov J.V.;
R7 "Ricin";
R7 Toxicon 39:1723-1728(2001).
RN
R8] X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RA MEDLINE-87163593; PubMed-3585197;
R7 Montfort W., Villafraza J.E., Monzingo A.F., Ernst S.R., Katzin B.,
R7 Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;
R7 "The three-dimensional structure of ricin at 2.8 A.";
RL J. Biol. Chem. 262:5398-5403(1987).
RN
R9] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RA MEDLINE-91352004; PubMed-1881881;
R7 Katzin B.J., Collins E.J., Robertus J.D.;
R7 "Structure of ricin A-chain at 2.5 A.";
RL Proteins 10:251-259(1991).
RN
R10] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
RA MEDLINE-91352005; PubMed-1881882;
R7 Rutenber E., Robertus J.D.;
R7 "Structure of ricin B-chain at 2.5-A resolution.";
RL Proteins 10:260-269(1991).
RN
R11] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
RA MEDLINE-95082010; PubMed-7990130;
R7 Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,
R7 Pauplett R.A.;
R7 "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";
RL J. Mol. Biol. 244:410-422(1994).
RN
R12] X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
RA MEDLINE-96374222; PubMed-8780513;
R7 Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,
R7 Molina-Sylth M.C., Robertus J.D.;
R7 "Structure and activity of an active site substitution of ricin A
R7 chain.";
RL Biochemistry 35:11098-11103(1996).
RN
R13] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RA MEDLINE-97240820; PubMed-9086280;
R7 Yan X., Hollis T., Sylth M., Day P., Monzingo A.F., Milne G.W.,
R7 Robertus J.D.;
R7 "Structure-based identification of a ricin inhibitor.";
RL J. Mol. Biol. 266:1043-1049(1997).
RN
R14] MUTAGENESIS.
RA MEDLINE-93165632; PubMed-1287657;
R7 Kin Y., Robertus J.D.;
R7 "Analysis of several key active site residues of ricin A chain by
R7 mutagenesis and X-ray crystallography.";
RL Protein Eng. 5:775-779(1992).
CC
C1-] FUNCTION: Ricin is highly toxic to animal cells and to a less

CC extent to plant cells. The A chain is responsible for inhibiting
CC protein synthesis through the catalytic inactivation of 60S
CC ribosomal subunits. It acts as a glycosidase that removes a
CC specific adenine residue from an exposed loop of 28S ribosomal
CC RNA. As this loop is involved in the binding of elongation
CC factors, the modified ribosomes are unable to support protein
CC synthesis. The A chain can inactivate a few thousand ribosomes
CC per minute, thus inactivating them faster than the cell can make
CC new ones. A single A-chain molecule can therefore kill an animal
CC cell. The B chain binds to cell receptors and facilitates the
CC entry into the cell of the A chain; B chains are also responsible
CC for cell agglutination (lectin activity). It binds to beta-D-
CC galactopyranoside moieties.
CC
CC -] CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC
CC -] SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC
CC -] DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC
CC -] PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE
CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271
CC
CC -] SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 KIP SUPERFAMILY.
CC
CC -] CAUTION: CONTRAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb.slb.ch/announce/>
CC or send an email to license@slb.slb.ch).
CC
CC
CC EMBL: X03179; CAA26339.1; -;
CC EMBL: X52908; CAA37095.1; -;
CC EMBL: X02388; CAA26230.1; -;
CC EMBL: A12892; CAA01058.1; -;
CC PIR: A24041; RLCSD.
CC PDB: 2AA1; 31-JAN-94.
CC PDB: 1APG; 31-JAN-94.
CC PDB: 1FMP; 31-OCT-93.
CC PDB: 1IFS; 14-JAN-98.
CC PDB: 1IFT; 14-JAN-98.
CC PDB: 1IFU; 14-JAN-98.
CC PDB: 1R7C; 31-OCT-93.
CC PDB: 1OBS; 16-JUN-97.
CC PDB: 1OBT; 16-JUN-97.
CC PDB: 1BR5; 02-SEP-98.
CC PDB: 1BR6; 02-SEP-98.
CC GlycosultedB: P02879; -;
CC InterPro: IPR001574; RIP.
CC InterPro: IPR000772; Ricin_B_lectin.
CC Pfam: PF00161; RIP. 1.
CC Pfam: PF00652; Ricin_B_lectin; 6.
CC PRINTS: PR00396; SHIGARICIN.
CC SMART: SM00438; RICIN; 2.
CC PROSITE: PS00231; RICIN_B_LECTIN; 2.
CC PROSITE: PS00275; SHIGA_RICIN; 1.
CC Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
CC Glycoprotein; Lectin; Signal; 3D-structure.
CC
CC SIGNAL 1 35
CC CHAIN 36 302 RICIN A CHAIN.
CC PEPTIDE 303 314 LINKER PEPTIDE.
CC CHAIN 315 576 RICIN B CHAIN.
CC DOMAIN 321 448 RICIN B-TYPE LECTIN 1.
CC DOMAIN 451 575 RICIN B-TYPE LECTIN 2.
CC REPEAT 331 373 1-ALPHA.
CC REPEAT 374 414 1-BETA.
CC REPEAT 417 449 1-GAMMA.
CC REPEAT 462 497 2-ALPHA.
CC REPEAT 501 540 2-BETA.
CC REPEAT 543 570 2-GAMMA.

```

FT ACT SITE 212 212
FT DISULFID 294 316 INTERCHAIN.
FT DISULFID 334 353
FT DISULFID 377 394
FT DISULFID 465 478
FT DISULFID 504 521
FT CARBOHYD 45 45
FT CARBOHYD 271 271
FT CARBOHYD 409 409
FT CARBOHYD 449 449
FT CONFLICT 76 76
FT CONFLICT 551 551
FT STRAND 43 47
FT TURN 49 50
FT HELIX 53 67
FT STRAND 73 74
FT TURN 75 76
FT STRAND 77 79
FT TURN 88 90

Query Match 29.9%; Score 425.5; DB 1; Length 576;
Best Local Similarity 40.6%; Pred. No. 1,7e-30;
Matches 117; Conservative 45; Mismatches 103; Indels 23; Gaps 11;

OY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPIL-RQSTIPVSDAORFVLYVELTNOGGSITTAI 61
DB 52 TVQSYTNPIRVAGRLTGADYRHEIPVLPNRYGLPIN-QRFLIVELSHHAEISVTIAL 109
OY 62 DVTNLYVAYAGDQGYFLDAPDAE-----RHLEFGR-TRSLPPTGYDLEKXGH 115
DB 110 DVTNLYVAYAGDQGYFLDAPDAE-----RHLEFGR-TRSLPPTGYDLEKXGH 167
OY 116 RDQIPGIEELIQVSALRY---PGSGTRAQARSLLILQIMISEAARENPWFARROYI 171
DB 168 LRENIIEGNGPLEAISALYRSTGTQDPLTARSFICIQIMISEAARENPWFARROYI 227
OY 172 NSGESFLPDYMLEETSMGOOSTOVQSTGVENPFRLGISTGNEFTLSNRYDIASL 231
DB 228 RYNRSRAPSDDSVITLSEMSGRSLTAIOESNOGAFASPIQORRNGSFYDVSTILPII 287
OY 232 AIMLFVGRDRSSSDVRYMPLIRPLEYSGAVDVCYTASEPVRIV 279
DB 268 ALMYRCAPRPSO-----FSLIRPVYPMFN--DV-CMDEPIVRLV 328

RESULT 6
AGGL_RICCO STANDARD; PRT; 564 AA.
AC P06750;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Agglutinin precursor (RCA) [contains: Agglutinin A chain (rRNA N-
glycosidase) (RC 3.2.2.22); Agglutinin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86059449; PubMed=2999130;
RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;
RT "The primary sequence of Ricinus communis agglutinin. Comparison with
RT ricin."
RU J. Biol. Chem. 260:15682-15686(1985).
RN [2]
RP SEQUENCE OF 303-564.
RC Tissue=Seed;
RA Arai T., Yoshioka Y., Funatsu G.;
```

```

RT "The complete amino acid sequence of the B-chain of the Ricinus
RT communis agglutinin isolated from large-grain castor bean seeds.";
RT Biochim. Biophys. Acta 872:277-285(1986).
RN [3]
RP SEQUENCE OF 303-337.
RX MEDLINE=80178723; PubMed=6768555;
RA Lin T.T.-S., Li S.S.-L.;
RT "Purification and physicochemical properties of ricins and
RT agglutinins from Ricinus communis.";
RL Eur. J. Biochem. 105:453-459(1980).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12089; AAA33869.1; -
DR EMBL; S40368; AAB22584.1; -
DR PIR; A24261; RLC5AG.
DR HSPD; P02879; IBR6.
DR GlycoStatedB; P06750; -.
DR InterPro; IPR001574; RIP.
DR InterPro; IPR000772; Ricin_B_lectin.
DR pfam; PR00161; RIP: 1.
DR pfam; PR00652; Ricin_B_lectin: 6.
DR SMART; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LLECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 290 POTENTIAL.
FT PROPEP 291 302 AGGLUTININ A CHAIN.
FT CHAIN 303 564 LINKER PEPTIDE.
FT DOMAIN 309 436 AGGLUTININ B CHAIN.
FT DOMAIN 439 563 RICIN B-TYPE LECTIN 1.
FT REPEAT 319 361 RICIN B-TYPE LECTIN 2.
FT REPEAT 362 402 1-ALPHA.
FT REPEAT 403 421 1-BETA.
FT REPEAT 422 437 1-GAMMA.
FT REPEAT 438 455 2-ALPHA.
FT REPEAT 456 489 2-BETA.
FT REPEAT 490 528 2-GAMMA.
FT REPEAT 529 558 2-GAMMA.
FT ACT_SITE 200 200 BY SIMILARITY.
FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
FT DISULFID 322 341 BY SIMILARITY.
FT DISULFID 365 382 BY SIMILARITY.
FT DISULFID 453 466 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 331 331 F -> T (IN REF. 2).
FT CONFLICT 362 362 N -> D (IN REF. 2).
FT CONFLICT 374 374 R -> G (IN REF. 2).
FT CONFLICT 404 404 R -> T (IN REF. 2).
FT CONFLICT 552 552 F -> V (IN REF. 2).
SQ SEQUENCE 564 AA; 62851 MW; D455F2AV2F60959 CRC64;

Query Match 29.7%; Score 423; DB 1; Length 564;
Best Local Similarity 39.7%; Pred. No. 2.8e-30;
Matches 114; Conservative 45; Mismatches 106; Indels 22; Gaps 10;
```

OY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPIL-RQSTIPVSDAORFVLYVELTNOGGSITTAI 61

DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).

OS Bryonia dioica (red bryony); Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; eudicotyledons; core eudicots; Rosidae; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eucosids I; Cucurbitales; Cucurbitaceae; Bryonia.

NCBI_TaxID=3652;

OX NCBI_TaxID=3652;

RN SEQUENCE FROM N.A.

RA Siegal C.B., Gawlak S.L., Marguardt H.;

RT "Bryodin 2, a ribosome-inactivating protein isolated from the plant Bryonia dioica.";

RL Patent number US5597569, 28-JAN-1997.

RP [2]

RP SEQUENCE OF 22-42.

RC TISSUE-Root;

RC MEDLINE=95151812; Pubmed=7849072;

RX Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marguardt H.;

RA "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates.";

RT Bioconj. Chem. 5:423-429(1994).

RL -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: I34238; NOT_ANNOTATED_CDS.

DR HSSP: P09989; IMRJ.

DR InterPro: IPR001574; RIP.

DR Pfam: PF00161; RIP: 1.

DR PRINTS: PR00396; SHIGARICIN.

DR PROSITE: PS00275; SHIGA_RICIN: 1.

KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.

KW Multigene family; Glycoprotein; Signal.

FT SIGNAL 1 21

FT CHAIN 22 282

FT ACT_SITE 183 183

FT CARBOHYD 25 25

FT SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

SO QUERY MATCH 24.8%; Score 354; DB 1; Length 282;

Best Local Similarity 39.3%; Pred. No. 1.7e-24;

Matches 95; Conservative 37; Mismatches 92; Indels 18; Gaps 9;

QY 4 TGDYFRFTLLRDVYSSGSFSENIPLRQSTIPVSDAQRVLVELTNOGGDSTIAAD 62

DB 32 TGAFTYFTRNRTKLTGCTPRVVDIPLVLAAGLA---RFOVLTLTNGSVYVALD 88

QY 63 VTNLYVAYOAGDQSYFLADAPDGAERHLFTGTRSSLPFGTSYDLEAYAGH--RDQIP 120

DB 89 VVAVVAVYAVRAGTAFELADASTEANNVLFAGINHVRLPYGNGDGLGLEYAAGRISRENI 148

QY 121 LGIEELIQSVASL--RYPGSTRQAQRSLITLIOMISEAFARNPIFMRARQYINGSESL 178

DB 149 LGSEISSAIGNFRNPGTSV---PRATIVITQVSEAFKYLEQVSE--NVGTFRK 203

QY 179 PDMYMLELETSMGQOSTOV--QSTDGVENNPRRLGT--SGNFTVLTLSNVADYASLAIMLF 234

DB 204 PDPAFLSLQANAGSLSPQIDIAQTRGGEFARPELVETLVSNPTFTVNVN--SPVYKGTALL 262

QY 235 LF 236

DB 263 LY 264

1:

RESULT 9

RIP2_MOMBA STANDARD; PRT; 286 AA.

ID RIP2_MOMBA

AC P29339;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ribosome-inactivating protein momordin II precursor (rRNA N-glycosidase) (EC 3.2.2.22).

DE Momordica balsamina (Bitter melon) (Balsam pear).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.

NCBI_TaxID=3672;

OX NCBI_TaxID=3672;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Seed;

RC MEDLINE=93027170; Pubmed=1408771;

RX Orliga M., Better M.;

RA "Momordin II, a ribosome inactivating protein from Momordica balsamina, is homologous to other plant proteins.";

RT Nucleic Acids Res. 20:4662-4662(1992).

RL -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: 212175; CAA78166.1; -

DR PIR: S25560; S25560.

DR HSSP: P09989; IMRJ.

DR InterPro: IPR001574; RIP.

DR Pfam: PF00161; RIP: 1.

DR PRINTS: PR00396; SHIGARICIN.

DR PROSITE: PS00275; SHIGA_RICIN: 1.

KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.

KW Multigene family; Glycoprotein; Signal.

FT SIGNAL 1 23

FT CHAIN 24 286

FT ACT_SITE 181 181

FT SEQUENCE 286 AA; 32031 MW; 3B89FF1AE6B25986 CRC64;

SO QUERY MATCH 23.9%; Score 340.5; DB 1; Length 286;

Best Local Similarity 35.2%; Pred. No. 2.8e-23;

Matches 89; Conservative 53; Mismatches 92; Indels 19; Gaps 9;

QY 4 TGDYFRFTLLRDVYSSGSFSENIPLRQSTIPVSDAQRVLVELTNOGGDSTIAAD 60

DB 33 TAKYTFEFL--EDFRATLPFSHKYDILPLAST---ISDRRLILDLSVALETISVA 86

QY 61 IDVTNLYVAYOAGDQSYFLADAPDGAERHLFTGTRSSLPFGTSYDLEAYAGH--RDQ 118

DB 87 IDVTNLYVAYOAGDQSYFLADAPDGAERHLFTGTRSSLPFGTSYDLEAYAGH--RDQ 145

QY 119 IPGLIEELIQSVASLRYPGSTRQAQRSLITLIOMISEAFARNPIFMRARQYINGSESL 178

DB 146 IDGLPPLSSAITLTFYNAQSAFSA--LLVLQTTAEARKYTERHVAVYV--ATNFK 201

QY 179 PDMYMLELETSMGQOSTOV--QOSTDGVENNPRRLGT--SGNFTVLTLSNVADYASLAIMLF 236

DB 202 PMLAITSLEQWMSALSKOIFLAQNGGKFRNPVDLKRPTGGEFOVYTNDSDVYKGNITLL 261

QY 237 VCDRPSDDVRY 249
DB 262 L-NSRASTADENF 273

RESULT 10

ID RIPS_TRIKI STANDARD: PRT: 289 AA.

AC P09989;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
OS (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS)
OC Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
NCBI_Taxid=3677;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-Maximowicz;
RA MEDLINE=91153657; PubMed=1999291;
RT Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia
RT coli.";
RL Gene 97:267-272(1991).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-Maximowicz; TISSUE=leaf;
RA MEDLINE=90256790; PubMed=2341400;
RT Chow T., Feldman R.A., Lovett M., Platak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
RT type I ribosome-inactivating protein.";
RL J. Biol. Chem. 265:8670-8674(1990).
[3]
RP SEQUENCE OF 24-270.
RC STRAIN-Maximowicz; TISSUE=tuberous root;
RA MEDLINE=90256789; PubMed=2341399;
RT Collins E.J., Robertus J.D., Lopest M., Stone K.L., Williams K.R.,
RT Wu P., Hwang K., Platak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular
RT models for abrin A-chain and alpha-trichosanthin.";
RL J. Biol. Chem. 265:8665-8669(1990).
[4]
RP SEQUENCE OF 24-270.
RC TISSUE=tuberous root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
RA Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THEF): history, chemistry and
RT application.";
RL Pure Appl. Chem. 58:789-798(1986).
[5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RA MEDLINE=94344957; PubMed=8066085;
RT Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RL Proteins 19:4-13(1994).
[6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA MEDLINE=95344383; PubMed=7619070;
RT Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT dephasing mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298(1995).
[7]
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT IS
CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -1- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC -1- specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE I RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M34858; AAA34206.1;
DR EMBL: J05434; AAA34206.1;
DR PIR: J00003; RUT27.
DR PIR: J00566; J00566.
DR PIR: A36273; A36273.
DR PIR: A36274; A36274.
DR PDB: 1MRJ; 07-FEB-95.
DR PDB: 1MRK; 07-FEB-95.
DR PDB: 1TCS; 10-JUL-95.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Plant defense; antiviral; Erotein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 270
FT PROPEP 271 289
FT ACET SITE 183 183
FT CONFLICT 57 60
FT CONFLICT 82 84
FT CONFLICT 87 87
FT CONFLICT 92 92
FT CONFLICT 143 144
FT CONFLICT 144 144
FT CONFLICT 196 196
FT CONFLICT 215 216
FT CONFLICT 231 231
FT CONFLICT 234 234
FT CONFLICT 246 246
FT CONFLICT 247 247
SQ
Query Match 23.6%; Score 336; DB 1; Length 289;
Best Local Similarity 33.9%; Pred. NO. 7.2e-23;
Matches 93; Conservative 56; Mismatches 95; Indels 30; Gaps 9;

QY 4 TGDYFRFTLLRDYVSSGSENEIPLRQSTIPSDAORFVLELNGGDSITAIDV 63
DB 33 TSSSYGVFTSNRKALPNERKLYDIPLR--QVALHLITNVADEITVAIDV 89
QY 64 TNLVVAAYQAGDSYFLDA-PDGERHLPTGTRS-SLPFTGSYTDLERYAGH-RDQIP 120
DB 90 TNNYIMGYRAGDTSYFENASATEAKYVRKADAMKVTLYSGNRYELDTAGKIRENIP 149
QY 121 LGIEELIOSVSLARKYPGGSTRAQARSLIILQMSIENARFPMFARQYINSGESFLPD 180
DB 150 LGLPALDSAITTLTFYNNANSASA--LMVLQISSEAKRKFTEDQIGKRPD--TFLPS 205
QY 181 MYLELETSWGOQSTNOVO--QSTDGVPNNPFRGISTGNFVLTSLNRYVIAIAIMLEVC 238
DB 206 LALISLEMSWALSKEQIOIASTINGOFESPVLINONORVITITNDVAGVTSNIALLLN 265
QY 239 RDRPSSSVRYKPLVIRPVLENGAGVDVTCAS 272
DB 266 RNNMAAND-----DDVPMQOS 281

RESULT 11
ID RIPS_TRIKI STANDARD: PRT: 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

CC specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AB045560; BAB19677.1; -
 CC HSSP: P16094; 1AHC.
 CC InterPro: IPR001574; RIP.
 CC Pfam: PF00161; RIP.1.
 CC PRINTS: PR00396; SHIGARICIN.
 CC PROSITE: PS00275; SHIGA_RICIN.1.
 CC Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
 CC toxin; Signal.
 CC KW SIGNAL. 1
 CC FT CHAIN. 1 21 POTENTIAL.
 CC FT ACT_SITE. 22 286 POTATIVE RIBOSOME-INACTIVATING PROTEIN.
 CC FT ACT_SITE. 185 185 BY SIMILARITY.
 CC FT CARBOHYD. 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD. 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD. 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 286 AA; 31771 MW; 4EPD4966E604DA41 CRC64;
 CC -----
 CC Query Match 21.6%; Score 308.5; DB 1; Length 286;
 CC Best local Similarity 33.2%; Pred. No. 2.1e-20;
 CC Matches 78; Conservative 52; Mismatches 92; Indels 13; Gaps 7;
 CC -----
 CC QY 8 YFFRITLADYVSSGSFSENEIFLRQSTIPYSDAQRFLVELTNGSGTAAIDYNY 67
 CC DB 40 YSFTLSMRALPNAGDIYNIPLVPS--ISGSRITLMQSNBEGNTIMADYVNY 96
 CC QY 68 VVAVYAGDQSYFLRDA-PDGAERHLEFGTGRSSLPPTGTYDLERYA-GHRDQIPLGTEE 125
 CC DB 97 IMEYLVNGTSTYFENETDQALASKFEVQGTSTIPYSGNOKOVARERDSIPLGMA 156
 CC QY 126 LIOSVALRYPGSGTRAQARSILIIQIMISEAARFNIFWRAQYINS-GESFLPDWML 184
 CC DB 157 LQSAISLTLY--YDSRSAPVAFVLQITAEAAKYI--EKQIIDRISVSKVDLAAI 211
 CC QY 185 ELFTSGGQSTQVQ--QSDTGVENPNPRLGISTGNFVLTNSNVADYASLAIMLFV 237
 CC DB 212 SLENENSLSKDQIOLAKSNQGFQTPYKIIIDKGIILEVTNVSILVYTKNIMLL 266
 CC -----
 CC RESULT 15
 CC RIPL_BRYDI STANDARD: PRT; 290 AA.
 CC AC P3185; Q9S819;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 15-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)
 CC DE (EC 3.2.2.22) (BD1).
 CC OS Bryonia dioica (Red bryony).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC OC eucrosts I; Cucurbitales; Cucurbitaceae; Bryonia.
 CC OC NCBI_TaxId=3652;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 CC RC TISSUE=leaf;
 CC RX MEDLINE=97228081; PubMed=9115985;
 CC RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
 CC Siegal C.B.;
 CC RA "Molecular, biological, and preliminary structural analysis of
 CC recombinant bryodin I, a ribosome-inactivating protein from the plant
 CC Bryonia dioica".
 CC RT Biochemistry 36:3095-3103(1997).
 CC -----

RN [2]
 RP SEQUENCE FROM N.A.
 RA Siegal C.B.;
 RT "Cloning and expression of a gene encoding bryodin I from Bryonia
 RT dioica".
 RL Patent number US5541110, 30-JUL-1996.
 RN [3]
 RP SEQUENCE OF 24-66.
 RC TISSUE=seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecucchi P.C., Lazzerini A.M., Barbieri L., Stirpe F., Soria M.,
 RA Lappi D.;
 RT "N-terminal sequence of some ribosome-inactivating proteins".
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RN [4]
 RP SEQUENCE OF 24-43.
 RC TISSUE=ROOT;
 RX MEDLINE=95151812; PubMed=7849072;
 RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
 RA Marguardt H.;
 RT "Characterization of ribosome-inactivating proteins isolated from
 RT Bryonia dioica and their utility as carcinoma-reactive
 RT immunokonjugates".
 RL Bioconjug. Chem. 5:423-429(1994).
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.
 CC -1- CATALYTIC ACTIVITY: Endopolydactylase of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO
 CC PRODUCE A SHORTER PROTEIN.
 CC -1- BIOCHEMICAL: Especially useful as immunotoxin for
 CC pharmacological applications as it has low toxicity in rats and
 CC mice but is potent once inside target cells.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: J24020; -; NOT_ANNOTATED_CDS.
 CC PIR: S16491; S16491.
 CC PDB: 1BRV; 04-MAR-98.
 CC InterPro: IPR001574; RIP.
 CC Pfam: PF00161; RIP.1.
 CC PRINTS: PR00396; SHIGARICIN.
 CC PROSITE: PS00275; SHIGA_RICIN.1.
 CC Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
 CC KW 3D-structure; Multigene family; Glycoprotein; Signal.
 CC FT SIGNAL. 1 23
 CC FT CHAIN. 1 24
 CC FT PROPEP. 271 290 RIBOSOME-INACTIVATING PROTEIN BRYODIN I.
 CC FT ACT_SITE. 183 183 MISSING IN MATURE PROTEIN.
 CC FT ACT_SITE. 212 212 BY SIMILARITY.
 CC FT CARBOHYD. 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD. 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT MTTACEN. 212 212 E-K: REDUCES ACTIVITY 10-FOLD.
 CC FT MTTACEN. 61 65 RSSIS -> LRHXI (IN REF. 3).
 CC SQ SEQUENCE 290 AA; 31788 MW; E96CDD9C031AA2DB CRC64;
 CC -----
 CC Query Match 20.9%; Score 297.5; DB 1; Length 290;
 CC Best local Similarity 32.9%; Pred. No. 2.1e-19;
 CC Matches 82; Conservative 54; Mismatches 86; Indels 25; Gaps 9;
 CC -----
 CC QY 5 GDEYFR-----FTFLRDYVSSGSFSENEIFLRQSTIPYSDAQRFLVELTNG 53
 CC DB 23 GDSYFRLSGATTTGTVGPIKMLREALPYERKYNIPLRSS--ISGSGRYTLHLTLNYA 79
 CC QY 54 GDSITAIQVNTLVVVAQAGDQSYFLRDA-PDGAERHLEFGTGRSSLPPTGTYDLER 111
 CC -----

Thu Mar 27 07:07:03 2003

us-09-627-165e-16.rsp

Page 13

```

Db      80  DEISVANDVTNYINGVLADGVSYFFNENASTEAKRVFEDAKKVTPLPSYGNERTQCT 139
      112  YAGH-RDQIPGIEELIYOSVSAKRPQCGSTRQARQSLITLQIMISEARFNFTMARQY 170
      140  AAGKREINPIGIPALDSATITLVYYVASSASA--LLVLQSTASARVRYKITEQIGKR 197
QY      171  INSGSEIPPMYMLEETSMGQOSTQYO--QSTDGVEFNPPRLGISTGAFVLSN--VRD 226
      198  VD--KTEFLPSIATISLENNWSALSQIQIASJFNQGFSSPYVLIDGNGRVSITNASRV 255
QY      227  VTASLAIML 235
Db      256  VTSNIALLL 264

```

Search completed: March 18, 2003, 08:25:46
Job time : 15.6181 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:18 ; Search time 37.1533 Seconds
(without alignments)
1547.297 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425
Sequence: 1 HOTGDEYFRITLLRDYVS.....NSGAVDDVCTCTASEPTVRIV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	1417	99.4	565	10 Q8W243	Q8W243 viscum albu
2	1264.5	86.7	531	10 Q8RXH6	Q8RXH6 viscum albu
3	1187	83.3	249	10 Q8RXH7	Q8RXH7 viscum albu
4	545	38.2	580	10 Q94BW3	Q94BW3 cinamomum
5	536	37.6	549	10 Q94BW2	Q94BW2 cinamomum
6	535	37.6	580	10 Q94BW4	Q94BW4 cinamomum
7	535	37.5	581	10 Q94BW5	Q94BW5 cinamomum
8	481	33.8	528	10 Q06076	Q06076 abrus prec
9	480	33.7	252	10 Q08761	Q08761 abrus prec
10	474	33.3	251	10 Q08237	Q08237 abrus prec
11	473	33.2	251	10 Q08235	Q08235 abrus prec
12	467	33.8	251	10 Q08237	Q08237 abrus prec
13	458	32.1	252	10 Q08760	Q08760 abrus prec
14	428.5	30.1	592	10 Q082E7	Q082E7 iris hollan
15	427	30.0	570	10 Q41358	Q41358 sambucus ni
16	426.5	29.9	573	10 Q0W2E8	Q0W2E8 iris hollan

17	425	29.8	564	10 Q9AVR2	Q9AVR2 sambucus eb
18	422.5	29.6	541	10 Q41174	Q41174 rictinus com
19	422.5	29.6	565	10 Q04071	Q04071 sambucus ni
20	421	29.5	570	10 Q22415	Q22415 sambucus ni
21	416	29.2	547	10 Q9M6E9	Q9M6E9 abrus prec
22	403	28.3	563	10 Q04367	Q04367 sambucus ni
23	401.5	28.2	604	10 Q9M654	Q9M654 polyonatum
24	390	27.4	563	10 Q94552	Q94552 sambucus ni
25	387.5	27.2	569	10 P93543	P93543 sambucus ni
26	383	26.9	566	10 Q04072	Q04072 sambucus ni
27	356.5	25.0	603	10 Q9M653	Q9M653 polyonatum
28	338.5	23.8	286	10 Q9FUV7	Q9FUV7 memordica c
29	337	23.6	289	10 Q94KE4	Q94KE4 trichosan
30	333.5	23.4	286	10 Q41257	Q41257 memordica c
31	333.5	23.3	254	10 Q9P5H2	Q9P5H2 memordica c
32	329	23.1	289	10 Q41216	Q41216 memordica c
33	327.5	23.0	300	10 Q04357	Q04357 iris hollan
34	324	22.7	247	10 Q9LRE3	Q9LRE3 trichosan
35	319.5	22.4	300	10 Q04356	Q04356 iris hollan
36	303.5	21.3	270	10 Q41611	Q41611 trichosan
37	303	21.3	258	10 Q04358	Q04358 iris hollan
38	302.5	21.2	258	10 Q9S9E4	Q9S9E4 gelonlum mu
39	291.5	20.5	278	10 Q00980	Q00980 lufia cylin
40	290.5	20.4	293	10 Q8S452	Q8S452 latrophle cu
41	289.5	20.3	293	10 Q8VYU0	Q8VYU0 latrophle cu
42	263.5	18.5	305	10 Q8W4U4	Q8W4U4 bougaluvj11
43	227	15.9	365	10 Q8RY69	Q8RY69 gnostemma
44	225	15.9	382	10 Q8SA43	Q8SA43 abrus prec
45	212	14.9	302	10 P93077	P93077 clerodendru

ALIGNMENTS

RESULT 1	ID	Q8W243	PRELIMINARY:	PRT:	565 AA.
AC	Q8W243:				
DT	01-MAR-2002 (TREMUREL. 20, Created)				
DT	01-MAR-2002 (TREMUREL. 20, Last sequence update)				
DT	01-JUN-2002 (TREMUREL. 21, Last annotation update)				
DE	VCA precursor.				
OS	Viscum album subsp. coloratum.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
OC	Santalales; Viscaceae; Viscum.				
OX	NCBI_TaxID=159976;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Park W.-B., Lyu S.;				
RT	*Cloning of Viscum album subsp. coloratum (Korean mistletoe).*				
RL	Biochem. Biophys. Res. Commun. 0:0-0(2002).				
DR	EMBL; AF369961; AAI40417.1; ..				
DR	InterPro; IPR001899; Gram_pos_anchor.				
DR	InterPro; IPR000772; Ricin_B_lectin.				
DR	InterPro; IPR001574; RIP.				
DR	Pfam; PF00652; Ricin_B_lectin; 6.				
DR	Pfam; PF00161; RIP; 1.				
DR	PRINTS; PR00396; SHIGARICIN.				
DR	SMART; SM00458; RICIN; 2.				
DR	PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.				
DR	PROSITE; PS50231; RICIN_B_LECTIN; 2.				
KW	SIGNAL.				
FT	SIGNAL	1	22	POTENTIAL.	
FT	CHAIN	23	273	VCA ALPHA CHAIN.	
FT	CHAIN	309	565	VCA BETA CHAIN.	
SQ	SEQUENCE	565 AA; 62401 MW; 991E3994DA005F11 CRC64;			
Query Match		99.4%;	Score 1417;	DB 10;	Length 565;
Best local Similarity		99.6%;	Pred. No. 4.9e-121;		
Matches 278;		Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;	
OY	1	HOTGDEYFRITLLRDYVS	SGFSFNSNEIPLRQSTTPVSDAQRFLVELTNGGDSITAA	60	

```

Db 43 HQTGDEYFRFTLLRDVSSGFSNEIPLLRSTIPVSDAQRFLVELTNOGDSITAA 102
|||
Qy 61 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 120
|||
Db 103 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 162
|||
Qy 121 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 180
|||
Db 163 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 222
|||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 240
|||
Db 223 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 282
|||
Qy 241 RPSSSDRYWPLVIRPVLENSGAVDDVCTAASEPTVRIY 279
|||
Db 283 RPSSSDRYWPLVIRPVLENSGAVDDVCTAASEPTVRIY 321
|||

```

RESULT 2

```

ID 08RXH6 PRELIMINARY: PRT: 531 AA.
AC 08RXH6:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Lectin chain A isoform 1 (Fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_Taxid=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramasivam M., Mista V., Srinivasan A., Singh T.P.;
RT Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1 and
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081149; AAL87006.1; -.
FT NON_TER 1 249
FT CHAIN 1 249 LECTIN CHAIN A ISOFORM 1.
FT CHAIN 269 531 LECTIN B CHAIN.
SQ SEQUENCE 531 AA; 58802 MW; 18244BEEFE35422 CRC64;

```

Query Match 88.7%; Score 1264.5; DB 10; Length 531;
 Best Local Similarity 88.2%; Pred. No. 4e-107;
 Matches 246; Conservative 15; Mismatches 13; Indels 5; Gaps 1;

```

Qy 1 HQTGDEYFRFTLLRDVSSGFSNEIPLLRSTIPVSDAQRFLVELTNOGDSITAA 60
|||
Db 10 HQTGDEYFRFTLLRDVSSGFSNEIPLLRSTIPVSDAQRFLVELTNOGDSITAA 69
|||
Qy 61 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 120
|||
Db 70 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 129
|||
Qy 121 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 180
|||
Db 130 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 189
|||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 240
|||
Db 190 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 249
|||
Qy 241 RPSSSDRYWPLVIRPVLENSGAVDDVCTAASEPTVRIY 279
|||
Db 250 RPSSSDRYWPLVIRPVLENSGAVDDVCTAASEPTVRIY 283
|||

```

RESULT 3

08RXH7

```

ID 08RXH7 PRELIMINARY: PRT: 249 AA.
AC 08RXH7:
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lectin chain A isoform 2 (Fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_Taxid=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramasivam M., Srinivasan A., Singh T.P.;
RT Viscum Album (Indian) mRNA for Mistletoe Lectin Chain A, isoform 2.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081148; AAL87005.1; -.
FT NON_TER 1 249
FT CHAIN 1 249
SQ SEQUENCE 249 AA; 27944 MW; 89FAFB78309A83B3 CRC64;

```

Query Match 83.3%; Score 1187; DB 10; Length 249;
 Best Local Similarity 97.9%; Pred. No. 1.7e-100;
 Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 HQTGDEYFRFTLLRDVSSGFSNEIPLLRSTIPVSDAQRFLVELTNOGDSITAA 60
|||
Db 10 HQTGDEYFRFTLLRDVSSGFSNEIPLLRSTIPVSDAQRFLVELTNOGDSITAA 69
|||
Qy 61 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 120
|||
Db 70 IDVTNLYVAVOAGDQSYFLDAPDGAERHLFTGTRSSLPPTGYSYDLERYAGHRDIP 129
|||
Qy 121 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 180
|||
Db 130 LGTEELIOSVSLRYPGSGTFAORSLITLQMSAARFNPIFMRARQYINSGESFLPD 189
|||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 238
|||
Db 190 MYMLELTSWGOOSTOVQOSTDGVFNPNPRLGISTGNEVYTLNVRDYASLAIMLFVCRD 247
|||

```

RESULT 4

094BW3

```

ID 094BW3 PRELIMINARY: PRT: 580 AA.
AC 094BW3:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE rRNA - glycosylase precursor (EC 3.2.2.22).
OS Cinnaomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnaomum.
OX NCBI_Taxid=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Vang O., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnaomum proteins and study of their expression
RT patterns."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -| CATALYTIC ACTIVITY: ENDORHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RRNA.
CC -| SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039803; AAK82460.1; -.
DR InterPro: IPR000772; RICHN_B_Lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR Pfam: PF00161; RIP; 1.
DR PROSITE: PS0231; RICHN_B_Lectin; 2.
DR Hydrolase: Signal; 32
DR PROSITE: PS0231; RICHN_B_Lectin; 2.
DR Hydrolase: Signal; 32
KW SIGNAL 1
KW SIGNAL 32
KW POTENTIAL.

```

Thu Mar 27 07:07:04 2003

us-09-627-165e-16.rsp

Page 3

FT	CHAIN	33	580	TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT				CINNAMOMIN III.
SQ	SEQUENCE	580 AA:	64421 MW; 940010F01E7B558 CRC64;	
	Query Match		38.28; Score 545; DB 10; Length 580;	
	Best Local Similarity		47.44; Prcd. No. 3.1e-41;	
	Matches 137; Conservative		43; Mismatches 93; Indels 16; Gaps	
QY	4	TGDEFFRITLLNDYVSSGSFSENEPILRQ-STIPYSAORFVAVETLNGGDS-ITPAI 61		
		: : : : : : : : : : : : : : : : : : : :		
Db	44	TKTSYQTEALRQAQASGEHPGIPMRBSYVP--DSKRFILVELSNMADSPYTLAV 101		
QY	62	DYTNLYVAAQAQDSQYFLR-DAPDGAERLFTGTRSSLPFTSYTDLERAG-HRQI 119		
		: : : : : : : : : : : :		
Db	102	DYTNLYVAAKRTQSGSEFLRENPDPVATLMLPDYTKRTPFSSGYTDLERAVGERREI 161		
QY	120	PGIEELIQSVASLRYPG-GSTRADAQARSLLIILOMISEARFNPIEMWARYINSGESFL 178		
		: : : : : : : : : : : : : :		
Db	162	LIGMDPLEMNAISALMTSNLNOQRLAASLLIVTQMWAEVAFREIYEVASEITPAEMFR 221		
QY	179	PMYALLETSMQOOSTQYOGST-GVFNPFELGISTGNVYLSNRY-VIASLAILMF 236		
		: : : : : : : : : : : : : :		
Db	222	POPAPLSLENNKMSALNAVQSGNGGVFSSPVELASINMKRYVYGSVDRVLSGLATMF 261		
QY	237	VGR--DRSSSDRYRPLVIRVLEN-----SGANDVYTCASPYRI 278		
		: : : : : : : : : : : : : : : : : :		
Db	282	ICRSTRASSDQGITDLMLIRFLADVAADVAITADDDTCAADPEYRI 330		

RESULT 5

ID	Q9FV22	PRELIMINARY;	PRT;	549 AA.
AC	Q9FV22;			
DT	01-MAR-2001 (TRENBLER, 16, Created)			
DT	01-MAR-2001 (TRENBLER, 16, Last sequence update)			
DT	01-MAR-2001 (TRENBLER, 20, Last annotation update)			
DE	rRNA - glycosylase (EC 3.2.2.22) (Trenblert).			
OS	Cinnamomum camphora (Camphor trees).			
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; Magnoliopsida; Laurales; Lauraceae; Cinnamomum.			
OX	NCBI_TaxID=13429;			
NP	(1)			
RA	SEQUENCE FROM N.A.			
RA	Xie L., Liu W.-Y., Wang E.-D.;			
RT	Molecular cloning of cinnamomum A-, B-chain and the expression,			
RT	purification, characterization and mutagenesis of the A-Chain.",			
RL	Submitted (SEP-2000) to the EMBL/Genbank/DBD databases.			
CC	-1- CATALYTIC ACTIVITY: ENDOXYLYSTIS OF THE N-GLYCOSID BOND AT ONE			
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.			
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.			

Protein	Accession	Length	Score	DB	Length
DR EMBL: AE253548	AA68918.2	37.68	536	10	549
DR HSSP: P02879	2A1	46.78	536	10	549
DR InterPro: IPR000772	Ricin_B_lectin	46.78	536	10	549
DR InterPro: IPR001574	RIP	46.78	536	10	549
DR Pfam: PF00682	Ricin_B_lectin	46.78	536	10	549
DR Pfam: PF00161	RIP	46.78	536	10	549
DR PRINTS: PR00396	SHIGARICIN	46.78	536	10	549
DR SMART: SM00458	RICIN	46.78	536	10	549
DR PROSITE: PS50231	RICIN_B_LECTIN	46.78	536	10	549
KW Hydrolase; Toxin		46.78	536	10	549
FT NON_TER		46.78	536	10	549
SEQUENCE	549 AA; 60648 MW; 02607FEE607CA44B0 CRC64;	46.78	536	10	549

QY 2 QTTGGEEVREFITLNDYVSSGGSFNPSEIPLLRQ-STIPVSDKQREVLVELTNPQGDS-ITYA 59
 :
 Db 10 KATSTSYQTGFALRQLASGEPEHGPVMKRSTVP--DKRRILVELSNMADAPSPYL 67
 :
 QY 60 AIDVTNLYVAHQAGQSIFEL-DAPDCAKHFTCTTRRSSLPPTSGYTLEIRYGA-HRD 117
 :

Db 68 ADVNNAYVAAYRTGSGQSPFLLEONNPPALIENTLPYTKRYTFPPFSSYTDLESGVAGERRE 127

Qy 118 QLELEIEELIQSVALRYRG--GSRPAQASLLITLIMSAPAFNPFLPMAQYINSGES 176

Db 128 ELLLOMDPLENAHSLAMISINLNOCPRLARSLIYIDMAVEAFRELEFYNGISIRREM 187

Qy 177 FLPPDNYMLELFTSGOOSTQVOVQSTID--GVNNPFLRGISTGNEVYTLNVRD--VYSLAIIM 233

Db 188 FRRDPAMLTSENNKMSLNAVQSGSGVESSPEVLRSTISNKPVYQSVSDRYISGLAIM 247

Qy 235 LEVGR--DRPSSDVRKPLVTRPYLEN-----SGADVDTCTASEPTVRI 278

Db 248 LPTCTSTDRASSDQFLIDHMLMLRLPLVDVAEVAITDMDNDTCCAPEPTVRI 298

RESULT 6

ID	PRELIMINARY:	PRT:	580 AA.
Q94BM4			
Q94BM4			
Q94BM4			
01-DEC-2001	(TrEMBLrel. 19, Created)		
01-DEC-2001	(TrEMBLrel. 19, Last sequence update)		
01-MAR-2002	(TrEMBLrel. 20, Last annotation update)		
DE	RNA - glycosidase precursor (EC 3.2.2.22).		
OS	Cinnamomum camphora (Camphor tree).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
XC	Spematiophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.		
NCBI_TextID:13429;			

RM	(1)	SEQUENCE FROM N.A.	
RP		Yang Q., Gong Z.Z., Liu W.Y.,	
RA		"Molecular cloning of three type 2 RIP (ribosome-inactivating protein)	
RT		genes encoding cinnamomin proteins and study of their expression	
RT		patterns";	
RC		Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.	
CC		- I. CATALYTIC ACTIVITY: ENDOHYDROLYTIC OF THE N-GLYCOSIDIC BOND AT ONE	
CC		SPECIFIC APEONINSE. ON THE 28S RRNA.	
CC		- II. SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.	
DR	EMBL:	AY039802; AKK8459.1; "	
DR	InterPro:	IPR000772; R1c1n_B.lectin.	
DR	InterPro:	IPR001574; RIP.	
DR	Pfam:	PF00652; R1c1n_B.lectin; 6.	
DR	Pfam:	PF00161; RIP; 1.	
DR	PROSITE:	PS0231; R1C1N_B.LECTIN; 2.	
KM	hydrolyase:	Signal; Toxin.	
FT	SIGNAL		
FT		1 32	POTENTIAL.
FT		3 580	TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT			CINNAMOMIN II.
QO	SEQUENCE	580 AA: 64265 MW: 3784289ECCECCERF CRC64:	

Query Match	37.6%	Score 536	DB 10	Length 580
Best Local Similarity 47.1%				
Pred. No. 2.1e-40				
Matches 136	Conservative 42	Mismatches 95	Indels 16	Gaps 10
QY	4	TGDEFFRFTLLRLRYVSSGSFNSNEIPLRLQ--STPVSDAORFLVVELTNOGDS--ITAAI	61	
Db	44	TKTSYTGIEIIRLQALSGCEEPFGFVPMHDSYTP--DSKRFLELISNMADSSVLA	101	
QY	62	DVTNLYVVAAYAGDQSYFLR--DAPDGAERHFTGTTSSSLPFFGSYTDLERYAGH--RDOI	119	
Db	102	DVTNLYVVAAYRTGQSQSFLEEDNPDALENNLLPDKRTYTPFGSGSYTDLERYAGHREBI	161	
QY	120	PLGIELLQVSALRYPG--GSTRPAOARSLILIMISEARFNPITMRARQVINGSCSFL	178	
Db	162	LGLMPLLNALISALMTSNLMOCALALRSILYIMVNAEAVRFFRIYRRESITRAEFMR	221	
QY	179	PMYMLLEETSMGQOSTOVQOSTD--GVENNPFRKGISTGFVLLSVNRV--VLAALIMLF	236	
Db	222	PPAPALSLENKWSALSNAAVQSGSVFSSPEVELRSLSNRPVYGVGSORVISCIALIMLF	281	
QY	237	VCR--DRPSSDVAKMPVLRPLEN-----SGAVADVDTTASEPVR	278	
Db	282	ICRSDDRTSSQDFIDHLIMIRPLIVADVDTTADNDNDTCADEPVR	330	

RESULT 7

ID Q94BW5 PRELIMINARY; PRT; 581 AA.

AC Q94BW5; 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA -glycosidase precursor (EC 3.2.2.22).
 OS Cinnamomum camphora (Camphor tree).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: Laurales: Lauraceae; Cinnamomum.
 NCBI_TaxID=13429;

RA Yang Q., Gong Z.Z., Liu W.Y.;
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
 RT genes encoding cinnamomin proteins and study of their expression
 RT patterns";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: AY039801; AAK82458.1; -.
 DR InterPro: IPR000772; R1CIN_B_lectin.
 DR InterPro: IPR001574; R1CIN_B_lectin.
 DR Pfam: PF00652; R1CIN_B_lectin; 5.
 DR Pfam: PF00161; R1P; 1.
 DR PROSITE: PSS0231; R1CIN_B_lectin; 2.
 DR Hydroxylase; Signal; Toxin.
 FT CHAIN 1 32 POTENTIAL.
 FT SIGNAL 33 581 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
 CINNAMOMIN I.
 SQ SEQUENCE 581 AA; 64215 MW; 6E8F5FBFB3D196 CRC64;

Query Match 37.5%; Score 535; DB 10; Length 581;
 Best Local Similarity 47.1%; Pred. No. 2.5e-40;
 Matches 136; Conservative 43; Mismatches 94; Indels 16; Gaps 10;

QY 4 TGDEFFRITLDRYVSSGFSNEIPLRO-STIPVDAORFVLELNOGDS-ITAI 61
 DB 4 TKTSTQPIELRAQLASGEPRGIPVMEKSTYP--DSKRFLVELSNMADSVYTLAV 101
 QY 62 DVTNLYVAYOAGDSYFLR-DAPDCAERHFTGTSSLPFGSTTDLERYAG-HRQI 119
 DB 102 DVTNLYVAYARTGSQSEFLRDPALNLPDKRTPEFGSGSTIDLESGVAGERREI 161
 QY 120 PLGIBELIQSVALRYPG-GSTRAQARSLILLIOMISEARFNPFRARQYINSGESFL 178
 DB 162 LIGMDPLENAISALWISNLNOQARALRSLIYIOMVAEAVRPFIEYRVRGSIARAEMFR 221
 QY 179 PDWYMLELETISWGQSTQVQSTD-GVFNPFRLGISTGNFVTLNVRD-VIASLAIMLF 236
 DB 222 PDPAWLSLENNKWSALSNAYQSGQGVSPPELRLSINRPPVYVGSVDRVLSGLAIMLF 281
 QY 237 VCR--DRPSSDVRYWPIVIRPVLEN-----SGAVDVTCTASEPTVRI 278
 DB 282 ICRSDIRASSDQFDIMIMIRPIYLVAVATDADNDCTADDEPPIVRI 330

RESULT 8

ID Q06076 PRELIMINARY; PRT; 528 AA.

AC Q06076; 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NCBI_TaxID=3816;

RP SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=8421313;
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoabrin subunits determined by cDNA
 RT sequencing: conservation and significance";
 RL J. Mol. Biol. 229:263-267(1993).
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: M98346; AAA32626.1; -.
 DR HSP: P11140; IABR.
 DR InterPro: IPR000772; R1CIN_B_lectin.
 DR InterPro: IPR001574; R1P.
 DR Pfam: PF00652; R1CIN_B_lectin; 6.
 DR Pfam: PF00161; R1P; 1.
 DR PRINTS: PRO0396; SHIGARICIN.
 DR SMART: SM00458; R1CIN; 2.
 DR PROSITE: PSS0231; R1CIN_B_lectin; 2.
 DR PROSITE: PSS0275; SHIGARICIN; UNKNOWN_1.
 DR Hydroxylase; Toxin.
 FT NON_TER 1 1
 FT NON_TER 528 528
 SQ SEQUENCE 528 AA; 58870 MW; 62ED42FB8F60F8 CRC64;

Query Match 33.8%; Score 481; DB 10; Length 528;
 Best Local Similarity 43.7%; Pred. No. 1.9e-35;
 Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

QY 4 TGDEFFRITLDRYVSSGFSNEIPLRO-STIPVDAORFVLELNOGDSITAIIV 63
 DB 13 TTSQYKQFELRQL-TGGILHIDIVLPDP-TVEERNRITVELSERESTLEVIDV 70
 QY 64 TNLVYVAYOAGDSYFLRDAPDCAERHFTGTSSLPFGSTTDLERYAG-HRQIPLG 122
 DB 71 TNLVYVAYAGDSYFLRDAPDCAERHFTGTSSLPFGSTTDLERYAG-HRQIPLG 130
 QY 123 IEEILQSVSALRYPGSFRARQARSLILLIOMISEARFNPFRARQYINSGESFLPDMY 182
 DB 131 LQALTATIFLNSGASNDKARFLIYIOMISEARFNPFRARQYINSGESFLPDMY 190
 QY 183 MLELETSWGQSTQVQSTDGVF-NPFRLGISTGNFVTLNVRD-VIASLAIMLFRCGR 241
 DB 191 MLELETSWGQSTQVQSTDGVF-NPFRLGISTGNFVTLNVRD-VIASLAIMLFRCGR 249
 QY 242 PSSSDVRYWPIVIRPVLEN--SGAVDVTCTASEPTVRI 278
 DB 250 PNAQGS--PLIRSIRES---KICSSRYEPTVRI 279

RESULT 9

ID Q38761 PRELIMINARY; PRT; 252 AA.

AC Q38761; 096234;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NCBI_TaxID=3816;
 RN SEQUENCE FROM N.A.
 RA Eversen G., Mathiesen A., Sundan A.;
 RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE OF 2-252 FROM N.A.
 RC TISSUE=SEED;
 RX MEDLINE=94139756; PubMed=8307038;

RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 RT "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 RL Eur. J. Biochem. 219:83-87(1994).
 CC -1- CATALYTIC ACTIVITY: ENDOPHYCOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X54873; CA38655.1; -
 DR HSSP: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN.
 DR Glycosylase; Hydrolase; Toxin.
 KW NON_TER 252
 SQ SEQUENCE 252 AA; 28229 MW; 187BB84E134AEE5 CRC64;

Query Match 33.7%; Score 480; DB 10; Length 252;
 Best Local Similarity 44.1%; Pred. No. 8.2e-36;
 Matches 109; Conservative 36; Mismatches 78; Indels 24; Gaps 5;

OY 4 TGDEYFRFTLLRDVYSSGSFSEIPLRSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 14 TSQSYKQFIALERL-RGGLIHDIPLDPPT-TLQENRRITVELSNSDTSIEVGIDV 71
 OY 64 TNLVYVAYAGDSYFLRDAPDGAERHLEFGTTRSSLPFGSTYDLERYAGH-RDQIPLG 122
 DB 72 TNAIVYVAYAGDSYFLRDAPDGAERHLEFGTTRSSLPFGSTYDLERYAGH-RDQIPLG 131
 OY 123 IEELIOSVALRPGSGSTRAQASLTLLIOMISEARPNPFMRARQYINSGESFLPDMY 182
 DB 132 LQALTHGISFRRSGNDNEKARTLLVIIOYAAARRRYSINRNVSTGTATAPDPA 190
 OY 183 MLEETSGWGOSTVOQSTDGVFNPRFGISTGFTVTLNVWD-----VIASL 231
 DB 192 MISLENNMNLISRGVQESVODTFPNO-----VTLINIRNPIYDLSHPYAVL 241

OY 232 AMLFVC 238
 DB 242 AMLFVC 248

RESULT 10
 OY6236 PRELIMINARY; PRT; 251 AA.

AC OY6236;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 GN PCDNAC-1-E164A.
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreeae; Abrus.
 OX NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SEED;
 RX MEDLINE=94139756; PubMed=8307038;
 RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 RT "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 RL Eur. J. Biochem. 219:83-87(1994).
 CC -1- CATALYTIC ACTIVITY: ENDOPHYCOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X76721; CA54139.1; -
 DR HSSP: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.

KW Hydrolase; Toxin.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28040 MW; D57FCB182E0ECC9 CRC64;

Query Match 33.3%; Score 474; DB 10; Length 251;
 Best Local Similarity 43.7%; Pred. No. 2.9e-35;
 Matches 108; Conservative 36; Mismatches 79; Indels 24; Gaps 5;

OY 4 TGDEYFRFTLLRDVYSSGSFSEIPLRSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 13 TSQSYKQFIALERL-RGGLIHDIPLDPPT-TLQENRRITVELSNSDTSIEVGIDV 70
 OY 64 TNLVYVAYAGDSYFLRDAPDGAERHLEFGTTRSSLPFGSTYDLERYAGH-RDQIPLG 122
 DB 71 TNAIVYVAYAGDSYFLRDAPDGAERHLEFGTTRSSLPFGSTYDLERYAGH-RDQIPLG 130
 OY 123 IEELIOSVALRPGSGSTRAQASLTLLIOMISEARPNPFMRARQYINSGESFLPDMY 182
 DB 131 LQALTHGISFRRSGNDNEKARTLLVIIOYAAARRRYSINRNVSTGTATAPDPA 190
 OY 183 MLEETSGWGOSTVOQSTDGVFNPRFGISTGFTVTLNVWD-----VIASL 231
 DB 191 MISLENNMNLISRGVQESVODTFPNO-----VTLINIRNPIYDLSHPYAVL 240

OY 232 AMLFVC 238
 DB 241 AMLFVC 247

RESULT 11
 OY6237 PRELIMINARY; PRT; 251 AA.

AC OY6237;
 DT 01-FEB-1997 (TREMBlrel. 02, Created)
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreeae; Abrus.
 OX NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SEED;
 RX MEDLINE=94139756; PubMed=8307038;
 RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 RT "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 RL Eur. J. Biochem. 219:83-87(1994).
 CC -1- CATALYTIC ACTIVITY: ENDOPHYCOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X76722; CA54140.1; -
 DR HSSP: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.
 KW Hydrolase; Toxin.
 KW NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28055 MW; 6F64755C3D9EAB79 CRC64;

Query Match 33.2%; Score 473; DB 10; Length 251;
 Best Local Similarity 43.7%; Pred. No. 3.6e-35;
 Matches 108; Conservative 36; Mismatches 79; Indels 24; Gaps 5;

OY 4 TGDEYFRFTLLRDVYSSGSFSEIPLRSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 13 TSQSYKQFIALERL-RGGLIHDIPLDPPT-TLQENRRITVELSNSDTSIEVGIDV 70
 OY 64 TNLVYVAYAGDSYFLRDAPDGAERHLEFGTTRSSLPFGSTYDLERYAGH-RDQIPLG 122

```

DB      71  TNAAYVAVRAGTOSTFLADAPSSASDYLFTGTDHSLPFTGYGDLERMAHQSQTIPG 130
      123  IEEILOSVALRTPGGSTRAQARSLIIOMISEARFPIPMARROYINSGESFLPDMY 182
      DB    131  LQALTHGISFRRSGNDNEKARLIVIIOMVAALFRISNRYSIGTGFQPDAA 190
      QY    183  MLEETSWGQOSTVOQOSTDGVFNPNFRLLSTGNFVLSNRD-----VIASL 231
      DB    191  MISTENNMWMLSRGVQESVODTFPNQ-----VLTINIRNEPVIWDSLSHPTVAVL 240
      QY    232  AIMLFVC 238
      DB    241  AIMLFVC 247

RESULT 12
096235  PRELIMINARY: PRT: 251 AA.
AC 096235:
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE RNA: glycosidase (EC 3.2.2.22) (Fragment).
OS Abris precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abris.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEED;
RX MEDLINE=94139756; PubMed=8307038;
RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
RT "Cloning and expression of three abrin A-chains and their mutants
RL derived by site-specific mutagenesis in Escherichia coli."
CC Bur. J. Biochem. 219:83-87(1994).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: X76644; CAA54092.1;
DR HSSP: P11440; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PRO0396; SHIGARICIN.
DR HydroLase: Toxin.
KW NON-TER
FT 1
FT 251
SQ SEQUENCE 251 AA; 27997 MW; 3B60351839AEPB7E CRC64;

Query Match 32.8%; Score 467; DB 10; Length 251;
Best Local Similarity 43.3%; Pred. No. 1, 3e-34;
Matches 107; Conservative 36; Mismatches 80; Indels 24; Gaps 5;

```

```

RESULT 13
038760  PRELIMINARY: PRT: 252 AA.
ID 038760:
AC 038760:
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Abris-E (RNA N-glycosidase) (EC 3.2.2.22) (Fragment).
OS Abris precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abris.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RX MEDLINE=91201329; PubMed=2016300;
RA Evensen G., Mathiesen A., Sundan A.;
RT "Direct molecular cloning and expression of two distinct abrin A-
RT chains."
RL J. Biol. Chem. 266:6848-6852(1991).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SUBUNIT: DISULFIDE-LINKED OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: TO OTHER BACTERIAL AND PLANT RIBOSOME-INACTIVATING
CC PROTEINS. BELONGS TO TYPE 2 RIP.
DR EMBL: X54872; CAA38654.1;
DR HSSP: P11440; IABR.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PRO0396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW HydroLase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.
FT CHAIN
FT 1
FT 252
FT ABRIN E, A CHAIN (BY SIMILARITY).
SQ SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;

Query Match 32.1%; Score 458; DB 10; Length 252;
Best Local Similarity 45.6%; Pred. No. 8, 5e-34;
Matches 108; Conservative 34; Mismatches 91; Indels 4; Gaps 4;

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 15:20:00 : Search time 2037.82 Seconds
(without alignments) 11953.474 Million cell updates/sec

Title: US-09-627-165E-15

Sequence: 1 catcagacgacgagggagcaga.....aaaccacgtgtgcacgtgta 837

Scoring table: IDENTITY:NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_pl:*
- 8: gb_pl:*
- 9: gb_pt:*
- 10: gb_to:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcg_hum:*
- 31: em_hcg_inv:*
- 32: em_hcg_other:*
- 33: em_hcg_mus:*
- 34: em_hcg_pla:*
- 35: em_hcg_rtd:*
- 36: em_hcg_mam:*
- 37: em_hcg_vrl:*
- 38: em_sy:*
- 39: em_hcg_hum:*
- 40: em_hcg_mus:*
- 41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	826.4	98.7	1873	AF369961 Viscum al
2	695.2	83.1	1923	AF58957 Sequence 1
3	695.2	83.1	1923	AR164305 Sequence
4	692	82.7	1596	AX019438 Sequence
5	688.8	82.3	1602	AX081149 Viscum al
6	683.8	81.7	1598	AX019435 Sequence
7	677.4	80.9	747	AX081148 Viscum al
8	668.4	79.9	752	AX139571 Sequence
9	668.4	79.9	752	AF508914
10	643.2	77.1	774	AF164303
11	644.2	77.0	756	AF09000
12	641	76.6	763	AX019436
13	640.2	75.6	762	AF0894
14	632.8	75.5	762	AX019439
15	631.6	75.5	762	AX019437
16	628.2	75.1	705	AR164307 Sequence
17	620.4	74.1	747	AF508915 Viscum al
18	614	73.4	768	AX019440
19	588.6	70.3	768	AX139575 Sequence
20	579.4	69.2	753	AF508916 Viscum al
21	522	62.4	522	AX139585
22	274.4	32.8	1596	AX019441
23	246	29.4	752	AX019442
24	246	29.4	752	AX019443
25	230	27.5	357	AX139583
26	137.4	16.4	2010	AX039803
27	135.8	16.2	1932	AF259548
28	135.8	16.2	2034	AX039801
29	134.2	16.0	2070	AX039802
30	116.2	13.9	756	AF164302
31	114.6	13.7	1584	AF164302
32	114.6	13.7	2743	AF164302
33	114	13.6	753	AF164302
34	114	13.6	756	AF164302
35	113.2	13.5	1584	AF164302
36	112.4	13.4	753	AF164302
37	112.4	13.4	753	AF164302
38	112.4	13.4	753	AF164302
39	111.6	13.3	804	AF164302
40	111.6	13.3	804	AF164302
41	110.8	13.2	753	AF164302
42	98.6	11.8	1581	AF164302
43	93.4	11.2	1873	AF164302
44	93	11.1	1858	AF164302
45	89.4	10.7	1892	AF164302

ALIGNMENTS

RESULT 1	AF369961	1873 bp	DNA	linear	PLN 12-DEC-2001
LOCUS	AF369961	1873 bp	DNA	linear	PLN 12-DEC-2001
DEFINITION	Viscum album subsp. coloratum VCA precursor, gene, complete cds.				
ACCESSION	AF369961				
VERSION	AF369961.1	GI:17529700			
KEYWORDS	Viscum album subsp. coloratum, coloratum VCA precursor, gene, complete cds.				
SOURCE	Viscum album subsp. coloratum, coloratum VCA precursor, gene, complete cds.				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Santalales; Viscaceae; Viscum.				
REFERENCE	1 (bases 1 to 1873)				
AUTHORS	Park, W.-B. and Lyu, S.				
TITLE	Cloning of Viscum album subsp. coloratum (Korean mistletoe)				

JOURNAL Biochem. Biophys. Res. Commun. (2002) In press
REFERENCE 2 (bases 1 to 1873)
AUTHORS Park, W.-B. and Iyu, S.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2001) Chemistry, Seoul Women's University, 126,
Kongnung-2dong, Nowon-gu, Seoul 139-774, Korea
FEATURES
source
 /organism="Viscum album subsp. coloratum"
 /sub_species="coloratum"
 /db_xref="taxon:159976"
 <138..>1835
 /Product="VCA precursor"
 138..1835
 /codon_start=1
 /Product="VCA precursor"
 /protein_id="FAL49417.1"
 /db_xref="GI:17529701"
 /translation="MNAFLASRRAMSCFLMGLVFGATYKATKRYERLRKYTHQ
TTGDEYFRTYILRDYVSSGSFNEIPILROSTIPVSDQREVLVETINQSDSTTA
IDVTLVYVAQADQSYFLRDAPEGRHLPFTGTRSSLPFTGTYDLERYAGRDQ
IPGLIEELIOSVLSALRYSGSTRAQARSILILQIMISEARENPIDWRARQYINSGES
FLPDYMLELSTWGOOSTVOYSTGVENNPFLRGISTENFVTLNSRDVDTASLAIM
LFVCHDRSSSDYRWPLYIRVLENSGAVDVCTISEPTVIVGRDGLADYADLAK
FYNGNPDIOLMFDPNOLMTIRDGTIGSGCLTITGYAGVYVIEFDCAVREYATL
WQIMNGNIIIPRSNLVLAASGSSTLTLYQTVQVYFLGQMLAGNDPAREVYIYGF
GRLCHEANGASVSEYCGSGKEMKALYGSIRPEKONODCLTSGDSVSTENIV
SCSAGSGORHEFTNEGILINLNGLVMDVAQSNPSLRILIIYPATGKNQMLPVP"
sig_peptide 138..203
mat_peptide 204..956
misc_feature /Product="VCA alpha chain"
 957..1061
 /Note="Region: linker peptide"
 1062..1832
 /Product="VCA beta chain"
BASE COUNT 456 a 458 c 492 g 467 t
ORIGIN
Query Match 98.7%; Score 826.4; DB 8; Length 1873;
Best Local Similarity 99.3%; Pred. No. 3,7e-237;
Matches 830; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CATGAGACGAGCGGCGACGATATATTCGGTTCATCAGCTTCTCGAGATTATGTCGA 60
 |||||
DB 264 CATCAACGACGGGCGACGAAATATTCGGTTCATCAGCTTCTCGAGATTATGTCGA 323
QY 61 AGCGAGCTTTTCCAAATGATNCCACTCTGCTGATGATGATCCCGTCGGGAT 120
 |||||
DB 324 AGCGGAGCTTTTCCAAATGATNCCACTCTGCTGATGATGATCCCGTCGGGAT 383
QY 121 GCGCAAGATTTGTGTGTGACTCACCATTCAGGGGGAGACTCGATCAGCGCC 180
 |||||
DB 384 GCGCAAGATTTGTGTGTGACTCACCATTCAGGGGGAGACTCGATCAGCGCC 443
QY 181 ATTCAGCTTACCAATCTGTACGTGTGCTTACCAAGCAGCGACCAATCTTCTTTG 240
 |||||
DB 444 ATTCAGCTTACCAATCTGTACGTGTGCTTACCAAGCAGCGACCAATCTTCTTTG 503
QY 241 GCGCAGCACCAGCGCGCGGAAGGATCTCTTACCGCGCACACAGATCTCTCTC 300
 |||||
DB 504 GCGCAGCACCAGCGCGCGGAAGGATCTCTTACCGCGCACACAGATCTCTCTC 563
QY 301 CCATTTCAGGAGCTAGACAGATCTGGAGGATAGCGCGGATATGGAGACAGATCCCT 360
 |||||
DB 564 CCATTTCAGGAGCTAGACAGATCTGGAGGATAGCGCGGATATGGAGACAGATCCCT 623
QY 361 CTGGGTATAGAGCACTCAATTCATTCGCTCTCGCGCTTCTTATTCAGGCGGACGAC 420
 |||||
DB 624 CTGGGTATAGAGCACTCAATTCATTCGCTCTCGCGCTTCTTATTCAGGCGGACGAC 683
QY 421 CGGGCCCAAGCTCTCTCTTAAATCTCAATTCAGATGATTCGAGGCGCGAGATTC 480
 |||||
DB 684 CGGGCCCAAGCTCTCTCTTAAATCTCAATTCAGATGATTCGAGGCGCGAGATTC 743

QY 481 AATCCATCTTTTGGAGGCTCGCCCATACATTAACAGGCGGAGATCTTCTCCGAC 540
 |||||
DB 744 AATCCATCTTTTGGAGGCTCGCCCATACATTAACAGGCGGAGATCTTCTCCGAC 803
QY 541 ATGTCATGCTCGAGCTGAGACACTAGTGGGCGCAACATCCAGCAGTCAGACT 600
 |||||
DB 804 ATGTCATGCTCGAGCTGAGACACTAGTGGGCGCAACATCCAGCAGTCAGACT 863
QY 601 ACGGATGCGCTTTTAAATACCATTTTCGGTGGTATATACCGGATTACTTGACG 660
 |||||
DB 864 ACGATGCGCTTTTAAATACCATTTTCGGTGGTATATACCGGATTACTTGACG 923
QY 661 TTGAGCATGCTGCGACGCTGATCGCGCTTACGATCATGTTGTTGATGTAAGGAC 720
 |||||
DB 924 TTGAGCATGCTGCGACGCTGATCGCGCTTACGATCATGTTGTTGATGTAAGGAC 983
QY 721 CGACCATCTTCGCGACGCTGATCGCGCTTACGATCATGTTGTTGATGTAAGGAC 780
 |||||
DB 984 CGACCATCTTCGCGACGCTGATCGCGCTTACGATCATGTTGTTGATGTAAGGAC 1043
QY 781 ACGGCGCGCTCGACGATGTTACTGTCAGTCTTCCGACCCAGCGTCGATCGT 836
 |||||
DB 1044 ACGGCGCGCTCGACGATGTTACTGTCAGTCTTCCGACCCAGCGTCGATCGT 1099
RESULT 2
LOCUS A58957 1923 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 1 from Patent WO9701636.
ACCESSION A58957
VERSION A58957.1 GI:3714428
KEYWORDS
SOURCE Viscum album.
ORGANISM Viscum album.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Santalales; Viscaceae; Viscum.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Lentzen, H., Eck, U., Baur, A. and Zinke, H.
TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: WO 9701636-A 1 16-JAN-1997;
MADDAUS AG KOELN (DE)
FEATURES
source
 /organism="Viscum album"
 /db_xref="taxon:3972"
 /tissue="LEAF"
 55..1749
 /codon_start=1
 /Product="PREPROMISTLETOE-LECTIN"
 /protein_id="CAA03513.1"
 /db_xref="GI:3714428"
 /translation="MNAVMDSRAMSCFLMGLVFGATYKATKRYERLRKYTHQ
TTGDEYFRTYILRDYVSSGSFNEIPILROSTIPVSDQREVLVETINQSDSTTA
IDVTLVYVAQADQSYFLRDAPEGRHLPFTGTRSSLPFTGTYDLERYAGRDQ
IPGLIEELIOSVLSALRYSGSTRAQARSILILQIMISEARENPIDWRARQYINSGES
FLPDYMLELSTWGOOSTVOYSTGVENNPFLRGISTENFVTLNSRDVDTASLAIM
LFVCHDRSSSDYRWPLYIRVLENSGAVDVCTISEPTVIVGRDGLADYADLAK
FYNGNPDIOLMFDPNOLMTIRDGTIGSGCLTITGYAGVYVIEFDCAVREYATL
WQIMNGNIIIPRSNLVLAASGSSTLTLYQTVQVYFLGQMLAGNDPAREVYIYGF
GRLCHEANGASVSEYCGSGKEMKALYGSIRPEKONODCLTSGDSVSTENIV
SCSAGSGORHEFTNEGILINLNGLVMDVAQSNPSLRILIIYPATGKNQMLPVP"
sig_peptide 55..153
mat_peptide 154..900
misc_feature /Product="A-CHAIN"
 901..957
 /Product="PROPEPTIDE"
 958..1746
 /Product="B-CHAIN"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN

Accession	AX019438	GI:10043377
Version	AX019438.1	
Keywords	synthetic construct.	
Source	synthetic construct.	
Organism	artificial sequences.	
Reference	1 (bases 1 to 1596).	
Authors	Welters, P., Stiefel, T., Voelter, M. and Morris, P.	
Title	Recombinant mistletoe lectins.	
Journal	Patent: WO 940109-A, 15.12.1993.	
	Welters, Peter (DE); Stiefel, Thomas (DE); Biotin Anzeimittel GmbH (DE); Voelter, Wolfgang (DE); Morris, Peter (GB)	
Location/Qualifiers		
Source	1..1596	
	/organism="synthetic construct"	
	/db_xref="taxon:32630"	
Base Count	387 a 418 c 419 g 372 t	
Origin		
Query Match	82.7%; Score 692; DB 6; Length 1596;	
Best Local Similarity	90.4%; Pred. No. 9.6e-197;	
Matches	756; Conservative 0; Mismatches 65; Indels 15; Gaps 1.	
1	CATCAACACACGCGGCGAGATATTTCCGGTTCATACAGCGCTCCGAGATATATCTCA	60
28	CATCAACACACGCGGCGAGAAATCTCCGGTTCATACAGCGCTCCGAGATATATCTCA	87
61	AGCGAAGGTTTCCATAGATAGATATACACTCTGCTCAGCTCAGCTCAGATCCCGTCCGAT	120
88	AGCGAAGGTTTCCATAGATAGATATACACTCTGCTCAGCTCAGCTCAGATCCCGTCCGAT	147
121	GCGCAAGATTTGTTGGTGGTGGTAACACACAACTCAGGCGGAGAGCTCAGTACAGCGCGCC	180
148	GCGCAAGATTTGTTGGTGGTGGTAACACACAACTCAGGCGGAGAGCTCAGTACAGCGCGCC	207
181	ATGACGCTTACCAATCTGTACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	240
208	ATGACGCTTACCAATCTGT	267
241	GCGGACGACACAGACGGCGCGGGAAGGACATCTCTTACACGCGACACACAGATCTCTCTC	300
268	GCGGACGACACAGACGGCGCGGGAAGGACACCTCTTACACGCGACACACAGATCTCTCTC	327
301	CCATTCACCGGAAGCTCAGATCTGTGAGCGATACGCGCGGATATAGGACACAGATCTCT	360
328	CCATTCACCGGAAGCTCAGATCTGTGAGCGATACGCGCGGATATAGGACACAGATCTCT	387
361	CTGGGATATGAGGAATCATATTCATCCGCTTGGCGGCTTGGTATATCAGGCGGACGACCC	420
388	CTGGGATATGAGGAATCATATTCATCCGCTTGGCGGCTTGGTATATCAGGCGGACGACCC	447
421	GCGGCGGACGCTGCTCCCTTATATATTCATTCATTCAGATATATTCAGGCGGCGGATTC	480
448	GCGGCGGACGCTGCTCCCTTATATATTCATTCATTCAGATATATTCAGGCGGCGGATTC	507
481	AATCCCATCTTTATGAGGCGCGCACATATTCATTCATTCAGATATATTCAGGCGGCGGATTC	540
508	AATCCCATCTTTATGAGGCGCGCACATATTCATTCATTCAGATATATTCAGGCGGCGGATTC	567
541	ATTCATATCTCTGACGCGGACGCTATTCATTCATTCAGATATATTCAGGCGGCGGATTC	600
568	GTTGATATCTCTGACGCGGACGCTATTCATTCATTCAGATATATTCAGGCGGCGGATTC	627
601	ACGGAATGGGCTTTTATTAACCATTTTCGTTGGGTATATTCACCGGTAACTTCTGTACG	660
628	ACGGAATGGGCTTTTATTAACCATTTTCGTTGGGTATATTCACCGCGGTAACTTCTGTACG	687
661	TTGAGCAATGTTGCGGACGTATGCGGACCTTACGATCATGTTGTTGATGTAGAGAC	720
688	TTGAGCAATGTTGCGGACGTATGCGGACCTTACGATCATGTTGTTGATGTAGAGAG	747
721	CGACCATCTCTCTGAGAGTGGCTTTTGGCGGCTGATACGACCGCTTGGAAAT	780

DB	748	CGCCATCTTCCTTGACGTCGACGTCATTTGGCCGCTGGTGTATACGACCCGCGAT-----	800
OY	781	AGCGGCGCGCGACGATGTTACCTACCTCTTCCGACACCGCTGCATGCT 836	
DB	801	-----AGCGATGATGTTACTCTGACGTCGCTCTTGACACTACGCTGCAGATTGT 848	
RESULT 5			
AY081149			
LOCUS	AY081149	1602 bp	mRNA
DEFINITION	Viscum album lectin chain A isoform 1 precursor, mRNA, partial cds.		
ACCESSION	AY081149		
VERSION	AY081149.1	GI:19526600	
KEYWORDS			
ORGANISM	Viscum album.		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Santalales; Viscaceae; Viscum.		
REFERENCE	1 (bases 1 to 1602)		
AUTHORS	Paramesivam, M., Mista, V., Srinivasan, A. and Singh, T. P.		
TITLE	Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1 and chain B		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1602)		
AUTHORS	Paramesivam, M., Mista, V., Srinivasan, A. and Singh, T. P.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-FEB-2002) Biophysics, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India		
FEATURES			
source	1..1602		
	/organism="Viscum album"		
	/db_xref="taxon:3972"		
	/tissue_type="leaf"		
	/country="India"		
	<1..1596		
	/note="toxic component"		
	/codon_start=1		
	/product="lectin chain A isoform 1 precursor"		
	/protein_id="AI167606.1"		
	/db_xref="GI:19526601"		
	/translation="TERKLRVRYHOTGEYRPRITLIRYVSSGSFSENEPILRQST IPVSDORFVLELNEIGSGDSTIRALIDVYNLVAIVAGQSTFLDARQRAETHLFT GTRSSLPENSGYDLERKGRKQILPILIDLOLQSTFLRPFQGNTRQAKSTLIL QMTSEARNPILYMKARQYINSGLASLMLFVCGEPSSSVRYKPLVIRPVADDTASAS EPTVYRGVYVMECDVNRDDFDHGNQIQMLPKSPSNQNLMLTKRDTIRNSCLT LYTATRGVYVMECDNDAVEATITWQENGTIINPSNITALASSGIRKGLTVVOTL DYLQGGMLAGNDTAPREVTYGVTFRDLCSNGSSQWVETCVSSQDNQMLAYDGGSI RPMQNDQCLTGRDSVSTVYINVSQASSQRMVPTNEGAILNLKRLADYKGSN PKRQIILYPAKGRQNMWLPV"		
	<1..747		
	/product="lectin chain A isoform 1"		
	/note="toxic component; ribosome-inactivating protein"		
	748..804		
	/note="Region: linker peptide"		
	805..1593		
	/product="lectin B chain"		
	1597..1602		
	3' UTR		
	BASE COUNT	394 a 415 c 419 g 374 t	
	ORIGIN		
Query Match	82.3%	Score 688.8	DB 8: Length 1602;
Best Local Similarity	90.2%	Pred. No. 8.8e-195;	
Matches 754;	Conservative 0;	Mismatches 67;	Indels 15; Gaps 1;
OY	1	CATCAGACGACGCGGCGACGAGATTTTCGGGTTCATCAGCGTTCTCGGAGATTAGTCTCA 60	
DB	28	CATCAACACCGCGGCGAGGATCTTCGCTTCATCAGCGTTCTCGGAGATTAGTCTCA 87	
OY	61	AGCGAGAGCTTTTCCATATGAGATACACTCTTGCGTCACTACGATCCCGCTCGCGAAT 120	
DB	88	AGCGAGAGCTTTTCCATATGAGATACACTCTTGCGTCACTACGATCCCGCTCGCGAAT 147	


```

OY 121 GCCCAAGATTTGTTGTTGTAAGTACCAATCAGGGGGAGACTGATCAGGCC 180
    |||||||
DB 148 GCCCAAGATTTGTTGTTGTTGTAAGTACCAATCAGGGGGAGACTGATCAGGCC 207
OY 181 ATCGAGTTACCAATCTGTATAGTGTGGCTTACCAAGCAGCGACCAATCTTATTTT 240
    |||||||
DB 208 ATCGAGTTACCAATCTGTATAGTGTGGCTTACCAAGCAGCGACCAATCTTATTTT 267
OY 241 CGGAGCGACAGCAGCGGGGAGAAAGCATCTTCCAGCGCACCAAGATCTCTTC 300
    |||||||
DB 268 CGGAGCGACAGCAGCGGGGAGAAAGCATCTTCCAGCGCACCAAGATCTCTTC 327
OY 301 CCATTACCGGAGAGTACAGAGATCTGAGAGATACGGCGGTCTATAGGACCAATCTCT 360
    |||||||
DB 328 CATTACAGAGAGTACAGAGATCTGAGAGATACGGCGGTCTATAGGACCAATCTCT 387
OY 361 CTGGTATAGAGAACTGATCAATCCGCTCGCGCTTGTATCCAGCGCGACAC 420
    |||||||
DB 388 CTGGTATAGAGAACTGATCAATCCGCTCGCGCTTGTATCCAGCGCGACAC 447
OY 421 CGGCGCAAGCTGTTCCCTTATATCTCATTCATTCAGATATCTCCAGCGCGAGATTC 480
    |||||||
DB 448 CATTACCAAGCTGTTCCCTTATATCTCATTCATTCAGATATCTCCAGCGCGAGATTC 507
OY 481 ATCCCATCTTTTGGAGGCTCGGCAATACATTAAACAGCGGGAGTCAATTTCTCCGAC 540
    |||||||
DB 508 ATCCCATCTTTTGGAGGCTCGGCAATACATTAAACAGTGGGCGTCAATTTCTCCGAC 567
OY 541 ATGATCATGCTGAGTGTGAGTGTGGGCGCAACAATCAGCGAAGTCCAGAGTCT 600
    |||||||
DB 568 GTGATCATGCTGAGTGTGAGTGTGGGCGCAACAATCAGCGAAGTCCAGAGTCT 627
OY 601 ACGATAGCGCTTTTATATACCCATTTCCGTTGGTATATCAACCGGTAACTTCTGTAGC 660
    |||||||
DB 628 ACCGAGGCGCTTTTATATACCCATTTCCGTTGGTATATCAACCGGTAACTTCTGTAGC 687
OY 661 TTGAGCAATGTTGCGGAGCTGATGCGACGCTTACGATATGTTTGTATAGAGAC 720
    |||||||
DB 688 TTGAGCAATGTTGCGGAGCTGATGCGACGCTTACGATATGTTTGTATAGAGAG 747
OY 721 CGACCATCTTCCCGAGCTGCGTATATGGCGCGTGTATACAGACCGCTTGTGAAAT 780
    |||||||
DB 748 CGGCGCATCTTCTGAGAGCTGCGTATATGGCGCGTGTATACAGACCGCTGAT 800
OY 781 AGCGGGCGCTGAGAGATGTTACCTGACGCTTCCGAAACCGACCGCGGACATCT 836
    |||||||
DB 801 -----AGCGGATGATGTTACTGCGAGTCTTCCGAAACCGTACGCGGATGT 848

RESULT 6
LOCUS AX019435 1598 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 12 from Patent WO9940109.
ACCESSION AX019435
VERSION AX019435.1 GI:10043374
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequences.
REFERENCE 1 (bases 1 to 1598)
AUTHORS Welter,P., Stiefel,T., Voelker,W. and Morris,P.
TITLE Recombinant mistletoe lectins
JOURNAL Patent: WO 9940109-A 12-12-AUG-1999;
WELTER PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH
(DE); VOELKER WOLFGANG (DE); MORRIS PETER (GB)
location/Owalfiers
FEATURES
source 1..1598
misc_feature 319..1322
misc_feature 373 a 396 c 392 g 356 t 81 others
BASE COUNT 373 a 396 c 392 g 356 t 81 others

```

```

ORIGIN
Query Match 81.7%; Score 683.8; DB 6; Length 1598;
Best Local Similarity 87.2%; Pred. No. 2,8e-194;
Matches 730; Conservative 37; Mismatches 54; Indels 16; Gaps 2;

OY 1 CATTACAGAGCGGCGAGCAATTTTCCGTTCAATACAGCTTCTCCGATATATGTCA 60
    |||||||
DB 28 CATTACAGAGCGGCGAGCAATTTTCCGTTCAATACAGCTTCTCCGATATATGTCA 87
OY 61 AGCGAGCTTCTCCATGTAGATACAGCTTGTGCTAGTACGTACAGATCCCGTCCGAT 120
    |||||||
DB 88 AGCGAGCTTCTCCATGTAGATACAGCTTGTGCTAGTACGTACAGATCCCGTCCGAT 147
OY 121 GCGCAAGATTTGTTGTTGTAAGTACCAATCAGGGGGAGACTGATCAGGCC 180
    |||||||
DB 148 GCGCAAGATTTGTTGTTGTTGTAAGTACCAATCAGGGGGAGACTGATCAGGCC 207
OY 181 ATCGAGTTACCAATCTGTATAGTGTGGCTTACCAAGCAGCGACCAATCTTATTTT 240
    |||||||
DB 208 ATCGAGTTACCAATCTGTATAGTGTGGCTTACCAAGCAGCGACCAATCTTATTTT 267
OY 241 CGGAGCGACAGCAGCGCGGAGAAAGCATCTTCCAGCGCACCAAGATCTCTCTCT 299
    |||||||
DB 268 CGGAGCGACAGCAGCGCGGAGAAAGCATCTTCCAGCGCACCAAGATCTCTCTCT 327
OY 300 CCCATTACCGGAGTGTACAGATCTGAGAGATACGGCGGTCTATAGGAGCAAGATCC 359
    |||||||
DB 328 CCATTCAATGGAAGTATACATCTGAGAGATACGGCGGTCTATAGGAGCAAGATCC 387
OY 360 TCGGGTATAGAGAACTGATCAATCCGCTCGCGGTCTGTTACAGGGGAGAGAC 419
    |||||||
DB 388 TCGGGTATAGAGAACTGATCAATCCGCTCGCGGTCTGTTACAGGGGAGAGAC 447
OY 420 CGGCGCAAGCTGTTCCCTTATATCTCATTCATTCAGATATCTCCAGCGCGAGATT 479
    |||||||
DB 448 GCGTATAGAGAACTGATCAATCCGCTCGCGGTCTGTTACAGGGGAGAGATT 507
OY 480 CATTCCATCTTTTGGAGGCTCGGCAATACATTAAACAGCGGGAGTCAATTTCTCCGA 539
    |||||||
DB 508 CATTCCATCTTTTGGAGGCTCGGCAATACATTAAACAGCGGGAGTCAATTTCTCCGA 567
OY 540 CATGTATCATGCTCGAGCTGAGAGCTGATGTTGGGCGCAACATTCAGCAGTCCAGATC 599
    |||||||
DB 568 CRTGTATCATGCTCGAGCTGAGAGCTGATGTTGGGCGCAACATTCAGCAGTCCAGATC 627
OY 600 TAGGATAGCGCTTTTATATACCCATTTCCGTTGGTATATCAACCGGTAACTTCTGTAGC 659
    |||||||
DB 628 TAGGATAGCGCTTTTATATACCCATTTCCGTTGGTATATCAACCGGTAACTTCTGTAGC 687
OY 660 GTTGAGCAATGTTGCGAGCTGATCCGAGCTTACGATCATGTTTGTATAGAGGA 719
    |||||||
DB 688 GTTGAGCAATGTTGCGAGCTGATCCGAGCTTACGATCATGTTTGTATAGAGGA 747
OY 720 CGGACCATCTTCCCGAGCTGCGTATATGGCGCGTGTATACAGACCGCTTGTGAAA 779
    |||||||
DB 748 GCGGCGCATCTTCTGAGAGCTGCGTATATGGCGCGTGTATACAGACCGCTGAT 801
OY 780 TAGCGCGCGCTGAGAGATGTTACCTGACGCTTCCGAAACCGACCGGATCT 836
    |||||||
DB 802 -----AGCGGATGATGTTACTGCGAGTCTTCCGAAACCGTACGCGGATGT 849

RESULT 7
LOCUS AY081148 747 bp mRNA linear PLN 17-MAR-2002
DEFINITION Viscum album lectin chain A isoform 2 precursor, mRNA, partial cds.
ACCESSION AY081148
VERSION AY081148.1 GI:19526598
KEYWORDS
SOURCE Viscum album.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 REFERENCE
 AUTHORS Paramasivam, M., Srihivasan, A. and Singh, T.P.
 TITLE Viscum Album (Indian) mRNA for Mistletoe Lectin Chain A, Isoform 2
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 747)
 AUTHORS Paramasivam, M., Srihivasan, A. and Singh, T.P.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2002) Biophysics, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India
 FEATURES
 source
 1..747
 /organism="Viscum album"
 /db_xref="taxon:3972"
 /tissue="leaf"
 /country="India"
 <1..>747
 /codon_start=1
 /product="lectin chain A isoform 2 precursor"
 /protein_id="AI87005.1"
 /db_xref="GI:19526599."
 /translation="YERLRURYHOTGDEYFRFTLLRDYVSGSPSELPILNOST
 IPVSDAQRFVVELINFGSDSYTAIDVNIYVAAGDQSPFLDADPDERLFT
 GTRSLPFTSYSDLERVAGHRDQIPGIEELIOSVALRYPGSGTRAAQSIILLI
 OMISEARFNPIFWVRVODINSSESLPDMYWELETSMSQOSTYQOSTDVFNNPF
 RLAISTGNFVTLISNVRDIASIALIMLFVCGE"
 <1..747
 /product="lectin chain A isoform 2"
 /note="toxic component; ribosome-inactivating protein"
 BASE COUNT 176 a 204 c 188 g 179 t
 ORIGIN
 mat_peptide
 Query Match 80.9% Score 677.4; DB 8; Length 747;
 Best Local Similarity 96.4%; Pred. No. 2, 2e-192;
 Matches 693; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

541 ATGTACATGCTCGAGCGAGACAGTAGTGGGGCCCAACATCCACGAAGTCCACAGTCT 600
 Db 568 ATGTACATGCTCGAGCGAGACAGTAGTGGGGCCCAACATCCACGAAGTCCACAGTCT 627
 Qy 601 ACGGATGGCGCTTTTATATACCATTTGCGGTATATACACGCGTAACTGTCAGC 660
 Db 628 ACGGATGGCGCTTTTATATACCATTTGCGGTATATACACGCGTAACTGTCAGC 687
 Qy 661 TTGACCAATGTGCGACAGTATGCCAGCTTACCATCATGTTGTTGTATAGGGA 719
 Db 688 TTGACCAATGTGCGACAGTATGCCAGCTTACCATCATGTTGTTGTATAGGGA 746
 RESULT 8
 AXI39571 762 bp DNA linear PAT 30-MAY-2001
 LOCUS AXI39571
 DEFINITION Sequence 1 from Patent EP1074560.
 ACCESSION AXI39571
 VERSION AXI39571.1 GI:14275207
 KEYWORDS
 SOURCE
 ORGANISM
 Viscum album subsp. coloratum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 REFERENCE
 AUTHORS Kang, J., Song, S., Suh, B., Lee, K., Do, M., Kwak, J., Song, B., Yoon, T.,
 Kim, J., and Park, C.
 TITLE Crude extract from Viscum album coloratum, and proteins and lectins
 isolated therefrom
 JOURNAL Patent: EP 1074560-A 1 07-FEB-2001;
 Mistle Biotech Co., Ltd. (KR)
 FEATURES
 source
 1..762
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 BASE COUNT 172 a 220 c 186 g 184 t
 ORIGIN
 Query Match 79.9% Score 668.4; DB 6; Length 762;
 Best Local Similarity 94.4%; Pred. No. 1, 1e-189;
 Matches 693; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Db 448 CCGGCCAAGCTGCTTCTTATCATCTCATTCAGATGATCTCCGAGGCCGCCAGATTC 507
QY 481 AATCCCATCTTTTGGAGGGCTCCCAATACATTACAGCGGGAGATCTTCTCCGAC 540
Db 508 AATCCCATCTTATGAGGGCTCCCAATACATTACAGCGGGAGATCTTCTCCGAC 567
QY 541 ATGACATGCTCGAGCTGAGACTAGTTGGGGCAACATCCAGCAATCCAGAGTCT 600
Db 568 AGGTACATCTTCAGCTGAGAGAGAGTGGGGCAACATCCAGCAATCCAGAGTCT 627
QY 601 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATACACCGGTACTCTGAGAC 660
Db 628 AGGATGGCGTTTATTAACCATTCGTTGGGTATATACACCGGTACTCTGAGAC 687
QY 661 TTGACCAATGTTCCGACGATGATCCGACCTTACGATCATGTTGTTGATAGGAC 720
Db 688 TTGACCAATGTTCCGACGATGATCCGACCTTACGATCATGTTGTTGATAGGAC 747
QY 721 CGACCATCTTCCTC 734
Db 748 CGGCATCTTCCTC 761

RESULT 9
AF508914 762 bp mRNA linear PLN 03-JUN-2002
LOCUS Viscum album subsp. coloratum lectin chain A isoform 1 mRNA,
DEFINITION partial cds.
ACCESSION AF508914
VERSION AF508914.1 GI:21314411
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
1 (bases 1 to 762)
Park, C.H., Lee, D.W., Kang, T.B., Lee, K.H., Yoon, T.J., Kim, J.B.,
Do, M.S., and Song, S.K.
CDNA cloning and sequence analysis of the lectin genes of the
Korean mistletoe (Viscum album coloratum)
Mol. Cells 12 (2), 215-220 (2001)
11710524
PUBMED 21566752
JOURNAL
MEDLINE
TITLE
AUTHORS
REFERENCE
PARK, C.H., LEE, D.W., KANG, T.B., LEE, K.H., YOON, T.J., KIM, J.B.,
DO, M.S., AND SONG, S.K.
DIRECT SUBMISSION
Submitted (02-MAY-2002) School of Bioscience and Food Technology,
Hanyang Univ., Haeundae-eup Namsong-ri, Pohang, Kyungbook 791-708,
Korea

FEATURES
source
1..762
Location/Qualifiers
/organism="Viscum album subsp. coloratum"
/sub_species="coloratum"
/db_xref="taxon:159976"
<1..>762
/note="lectin A1 chain; KML-A1"
/codon_start=1
/product="lectin chain A isoform 1"
/protein_id="AA046932.1"
/db_xref="GI:21314412"
/translation="YERLRIRVTHQHTGDEYFFITLLDYSVSSFSFNEPIILROST
IPVSDARVLELITNOCGDSITTAIDVTMLYVAIQADQSTFLRDAPDAERHFLT
GTRRSILPTFGSTIDLERAGRODITPLGRELIDSVSLRPGSNTRAKASHLILI
OMISEARNPILMKRQIISGGSFLPDVYILQLSTSGQSTOVYHSTQGVENPFI
RLTISTGVFTLSNVDVIALSLIMLFCEDRSSS"
misc_feature
748..>762
/note="Region: linker"
BASE COUNT 172 a 220 c 186 g 184 t
ORIGIN
Query Match 79.9%; Score 668.4; DB 8; Length 762;

Best Local Similarity 94.4%; Pred. No. 1.1e-189;
Matches 693; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 CATTAGACGAGCGGCGAGATATTTCCGGTTTCATCAGCTTCTCCGAGATTATGTCGA 60
Db 28 CATTAAACCCAGGGGCGAGCAATATTTCCGGTTTCATCAGCTTCTCCGAGATTATGTCGA 87
QY 61 AGCGAAGCTTTTCCATAGATAGATACACTCTGCGTACGTATACATCCCGTCCGAT 120
Db 88 AGCGAAGCTTTTCCATAGATAGATACACTCTGCGTACGTATACATCCCGTCCGAT 147
QY 121 GCGCAAAATTTGTTGTTGATGACATCCAGATACAGGGGAGATCATGACGCGCC 180
Db 148 GCGCAAAATTTGTTGTTGATGACATCCAGATACAGGGGAGATCATGACGCGCC 207
QY 181 ATGACCTTACCATCTGACCTGCTGCTTACCAAGCAGGAGCAATCTACTTTTG 240
Db 208 ATGACCTTACCATCTGACCTGCTGCTTACCAAGCAGGAGCAATCTACTTTTG 267
QY 241 CGGAGCAGCAGACGCGCGGAAAGGATCTCTTACCGGACAGCAGATCTCTCTC 300
Db 268 CGGAGCAGCAGACGCGCGGAAAGGATCTCTTACCGGACAGCAGATCTCTCTC 327
QY 301 CCATTACCGGAGCTACACAGATCTGGAGGATACCGGTCATAGGAGACAGATCCCT 360
Db 328 CCATTACCGGAGCTACACAGATCTGGAGGATACCGGTCATAGGAGACAGATCCCT 387
QY 361 CTGGCTTACAGCACTCATTCATCCGCTCGCGGCTCTTATCCAGGGGAGGAC 420
Db 388 CTGGCTTACAGCACTCATTCATCCGCTCGCGGCTCTTATCCAGGGGAGGAC 447
QY 421 CGGCCCCAGCTGCTCCCTTAAATCTCATTCAGTATGATGATCTCGAGGCGGAGATTC 480
Db 448 CGTCCCAACCTGCTCTCTTATTCATCTCATTCAGTATGATGATCTCGAGGCGGAGATTC 507
QY 481 AATCCCATCTTATGAGGGCTCGCCAAATACATTAAACAGGGGAGATCTTCTCCGAC 540
Db 508 AATCCCATCTTATGAGGGCTCGCCAAATACATTAAACAGGGGAGATCTTCTCCGAC 567
QY 541 ATGACATGCTCGAGCTGAGATGATGGGGCCCAACATCCAGCAATCCAGAGTCT 600
Db 568 ACGTACATCTTCAGCTGAGAGAGAGTGGGGCAACATCCAGCAATCCAGAGTCT 627
QY 601 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATACACCGGTACTCTGAGAC 660
Db 628 AGGATGGCGTTTATTAACCATTCGTTGGGTATATACACCGGTACTCTGAGAC 687
QY 661 TTGACCAATGTTCCGACGATGATCCGACCTTACGATCATGTTGTTGATAGGAC 720
Db 688 TTGACCAATGTTCCGACGATGATCCGACCTTACGATCATGTTGTTGATAGGAC 747
QY 721 CGACCATCTTCCTC 734
Db 748 CGGCATCTTCCTC 761

RESULT 10
AR164303 774 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 30 from patent. US 6271368.
DEFINITION AR164303
ACCESSION AR164303.1 GI:16235405
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 774)
AUTHORS Lentzen, H., Eck, J., Baur, A. and Zinke, H.
TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: US 6271368-A 30 07-AUG-2001;
FEATURES Location/Qualifiers
1..774
/organism="unknown"

BASE COUNT	173 a	232 c	185 g	184 t
ORIGIN				

AUTHORS Eck, J. and Schmidt, A.
TITLE RECOMBINANT FUSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS

Query Match	77.18;	Score 645.2;	DB 6;	Length 774;
Best local alignment	00.78	Score 1	DB 1	Length 100

JOURNAL
Patent: WO 9829540-A 7 09-JUL-1998;
INVENTOR: [REDACTED]
ATTORNEY: [REDACTED]

Matches 677; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

FEATURES	Location/Qualifiers
1	

QY 1 CATCAGACGAGGGCGACGATATTTCCGGTTCATCAGGCTTCTCCGAGATTATGCTCA 60

/organism="Viscum album"
/abundance="4000000000"

Db 34 CACCAGACCACGGTGAAGATATTTCCGGTTCATCAGGCTTCTCCGAGATTATGTCCTCA 93

```

/tissue_type="LEAF"

```

OY	601	ACGAGTGGCGTTTAAATTAACCATTGTGGCTGATTCACACGGGTACTGTCGAG	660
Db	628	ACGAGTGGCGTTTAAATTAACCATTGTGGCTGATTCACACGGGTACTGTCGAG	687
OY	661	TTGACCAATGTTCCGCACGTGATGCCAGCTTACCATCATTGTTTGTATATAGGAGAC	720
Db	688	TTGACCAATGTTCCGCACGTGATGCCAGCTTGGCGATCATGTTGTATATAGGAGAG	747
OY	721	CGACCATCT 729	
Db	748	CGACCATCT 756	
<hr/>			
RESULT 12			
LOCUS	AX019436	763 bp	DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 13 from Patent WO9940109.		
ACCESSION	AX019436		
VERSION	AX019436.1 GI:10043375		
KEYWORDS	.		
SOURCE	synthetic construct.		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 763)		
AUTHORS	Welters P., Stiefel T., Voelter W. and Morris P.		
TITLE	Recombinant mistletoe lectines		
JOURNAL	Patient NO 9940109-A 13.12.1999;		
	WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH		
	(DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)		
	Location/Qualifiers		
FEATURES	Source		
	1..763		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
BASE COUNT	misc_feature 163 a 215 c 175 g 170 t 40 others		
ORIGIN	319		
<hr/>			
Query Match			
Best Local Similarity 89.1%; Pred. No. 1.9e-181;			
Matches 655; Conservative 37; Mismatches 42; Indels 1; Gaps 1;			
OY	1	CATCACAGCAGCGGGGACGAATTTTCCGGTTTCATCAGCGTTCTCCGAGATTATGTCTCA	60
Db	28	CATCAACACGAGGGGAKAIAATCTTCGGTTTCATCAGCGTTCTCCGAGATTATGTCTCA	87
OY	61	AGCGGAAGCTTTTCCATGATGAGAATACACGTCCTGGTCACTGACATCCCGCTCGGAT	120
Db	88	AGCGGAAGCTTTTCCATGATGAGAATACACGTCCTGGTCACTGACATCCCGCTCGGAT	147
OY	121	GCCCAAGATTTTGTGTGTGTGGAATCTACCAATTCAGGGGGGAGATCGATCAGGCGCC	180
Db	148	GCCCAAGATTTTGTGTGTGTGGAATCTACCAACAGGGGGSRKACTCGATTACGGCGCC	207
OY	181	ATCGACGTTAACCATCTGTACGTGTGTGTGTACCAAGCAGCGACCAATCTACTTTTTG	240
Db	208	ATCGACGTTAACCATCTGTACGTGTGTGTGTACCAAGCAGCGACCAATCTACTTTTTG	267
OY	241	CGCGAGCAGCAGACGGCGGGGAAAAGGACTCTCTTACCGGACACAGATCGA-TCTCTCT	299
Db	268	CGCGAGCAGCAGACGGCGGGGAAAAGGACTCTCTTACCGGACACCGAATCTCTCT	327
OY	300	CCCATTCACCGGAAGCTACACAATGTGGAGGATTAAGCCGGTCATTAGGAGACAGATCCC	359
Db	328	CCCATTCACCGGAAGCTACMCYGATGTGSAGGATTAAGCCGGACATATAGGACAGATCCC	387
OY	360	TCGTGGTAAAGAGACATCATACGCTTCGGCGCTTCGTTATCAGCGGCGACAGC	419
Db	388	TCGTGGTAAAGAGACATCATATCGCTCKCGKCGCTTCGTTATCAGCGGCGGAGAC	447
OY	420	CGCGGCCAAGCTCGTCCCTTAATAAATCTCATTAATGATATCTCCGAGGCGCGAGATT	479
Db	448	GCCTGTCAGAGCTCGTCCCTTAATAAATCTCATTAATGATATCTCCGAGGCGCGAGATT	507

QY	480	CATCCCATCTTTGGAGGCGTCCGCAATACATTAACAGGCGGAGATCTTCTCCGCA	539
Db	508	CATCCCATCTTATATGAGAGKATCGCCAKAVATTAACAGTGGGMRCATTTCTCGCAGA <td>567</td>	567
QY	540	CATGTATAGTCGACGAGTGAAGATATGTGTGGGGCAACATACAGCAAGTCAGAGTC <td>599</td>	599
Db	568	CRGTTCACATCGTGAAGCTGAGACAGAGTGGGGCCCAATCCAGCAAGTCAGATTC <td>627</td>	627
QY	600	TACGATATGCGCTTTTATATACCATTTTCGGTGGATATTCACGGTAATCTGAGAC <td>659</td>	659
Db	628	AACCGATGCGCTTTTATATACCATTTTCGGTGGATATTCACGGTAATCTGAGAC <td>687</td>	687
QY	660	GTTTCACCATATGTTTCGACGATGATGCGACAGCTTAGCGATCATGTTGTTGTATAGGA <td>719</td>	719
Db	688	GTTTCACCATATGTTTCGACGATGATGCGACAGCTTAGCGATCATGTTGTTGTATAGGA <td>747</td>	747
QY	720	CGACCATCTCTCTC 734 <td></td>	
Db	748	CGCGCATCTCTCTC 762 <td></td>	
RESULT 13			
LOCUS	A90894	762 bp	DNA
DEFINITION	Sequence 1 from Patent WO9829540.		linear
ACCESSION	A90894		
VERSION	A90894.1		
KEYWORDS	GI:6739419		
SOURCE	European mistletoe.		
ORGANISM	Viscum album		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Sapnatales; Viscaceae; Viscum.		
REFERENCE	1 (bases 1 to 762).		
AUTHORS	Eck J and Schmidt A.		
TITLE	RECOMBINANT FUSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS		
JOURNAL	OP EUROPAN MISTLETOE VISCUM ALBUM		
	Patent: NO 9829540-A, 1.08-Jul-1998;		
FEATURES	ECK J DERGEN (DE); SCHMIDT ARNO (DE)		
source	Location/Qualifiers		
	1..762		
	/organism="Viscum album"		
	/db_xref="taxon:3572"		
	/tissue_type="LEAF"		
	4..>759		
	/function="TOXIC COMPONENT"		
	/note="FIG. 1A: NUKLEOTIDSEQUENZ UND ABGELEITETE AMINO		
	ACIDSEQUENZ"		
	/codon_start=1		
	/product="RECOMBINANT MISTLETOE LECTIN A-CHAIN"		
	/protein_id="CAB69393.1"		
	/db_xref="GI:6739420"		
	/translation="MYERISILTAHQTEGEFFRILLRDVYSSGFSNEIPILROS		
	TIPASDQRFVLELVETNGGDSILTAIDVTNLYVAYADGDSYFLRDAPEAEHLEF		
	TGTRRSFNGSYDLERYAGHRDIPGLIGDILLOSLATRPFGSGTQARSLILE		
	IQMSEARFNGILMARROYINGASGLPDVYMLLETLSMGDSSTOVSHSTDGCVNNPD		
	IRLAIIPGNVLTIVRDVIAIMLFCGRP"		
BASE COUNT	169 a	230 c	183 g 180 t
ORIGIN			
Query Match	76.5%:	Score 640.2:	DB 6: Length 762:
Best Local Similarity	92.7%:	Pred. No. 3.3e-181:	
Matches	672:	Conservative 0:	Mismatches 53: Indels 0: Gaps 0:
QY	1	CATCAGACACGCGGACGACGATATTCGCGGTATCATCAGCTTCTCGGATATATGATCTCA	60
Db	34	CACACACACACGCGGAGAGATATTCGCGGTATCATCAGCTTCTCGGATATATGATCTCA	93
QY	61	AGCGAGAGCTTTTCCATGAGATACCACTCTTGCGTCAAGTCACGATCCCGCTCGGAT	120
Db	94	AGCGAGAGCTTTTCCATGAGATACCACTCTTGCGTCAAGTCACGATCCCGCTCGGAT	153
QY	121	GCAGCAAGATTTGTGTGGTGGAGACTCACCAATCAGGGGGGAGAGATCTGATACGCGCCG	180

Db 154 GCGCAAGATTTGCTGTGTGAGCTCACCACACAGGGGAGAGCTCGATCCAGCGCC 213
|||
Oy 181 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGGGAGCAATCCACTTTTG 240
|||
Db 214 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGGGAGCAATCCACTTTTG 273
|||
Oy 241 GCGGACGACACAGCGGCGGAAAGCATCTCTTCAACCGGACACACAGATCCTCTC 300
|||
Db 274 GCGGACGACACAGCGGCGGAAAGCATCTCTTCAACCGGACACACAGATCCTCTC 333
|||
Oy 301 CCATTCACGGAAGCTACACAGATCTGAGCGATACCGGTCATAGGAGCAGATCCCT 360
|||
Db 334 CCATTCACGGAAGCTACACAGATCTGAGCGATACCGGATAGGAGCAGATCCCT 393
|||
Oy 361 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCTGTTATCCAGGCGGACAGCC 420
|||
Db 394 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCTGTTATCCAGGCGGACAGCC 453
|||
Oy 421 GCGGACGACGCTGCTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 480
|||
Db 454 CGTACCCAGCTGCTTCATTTTAACTCTCATTCAGATGATCTCGAGGCGCGAGATTC 513
|||
Oy 481 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGCGGAGATCATTTCTCCGAC 540
|||
Db 514 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGTGGGCGCTCATTTCTGCCAGAC 573
|||
Oy 541 ATGTACATGCTGAGCTGGAGACTAGTTGGGCGCCAAATCCAGCAATCCAGAGTCT 600
|||
Db 574 GTGTACATGCTGAGCTGGAGACTAGTTGGGCGCCAAATCCAGCAATCCAGAGTCT 633
|||
Oy 601 ACGATGAGGCTTTTAAATACCATTTGCGTGTGATTAACAGCGGATTAATCTGTCGAC 660
|||
Db 634 ACGATGAGGCTTTTAAATACCATTTGCGTGTGATTAACAGCGGATTAATCTGTCGAC 693
|||
Oy 661 TTGAGCAATGTCGACGCTGATCCGACGCTTACGATCTGTTGTATGAGGAC 720
|||
Db 694 TTGAGCAATGTCGACGCTGATCCGACGCTTACGATCTGTTGTATGAGGAG 753
|||
Oy 721 CGACC 725
|||
Db 754 CGACC 758
|||

RESULT 14
AX139573 762 bp DNA linear PAT 30-MAR-2001
LOCUS AX139573
DEFINITION Sequence 3 from Patent EP1074560.
ACCESSION AX139573
VERSION AX139573.1 GI:14275208
KEYWORDS
SOURCE
ORGANISM Viscum album subsp. coloratum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
1 (bases 1 to 762)
AUTHORS Kim,J., Song,S., Sub,B., Lee,K., Doo,M., Kwak,J., Song,B., Yoon,T.,
Kang,T. and Park,C.
TITLE Crude extract from Viscum album coloratum, and proteins and lectins
isolated therefrom
JOURNAL Patent: EP 1074560-A 3 07-FEB-2001;
Mistle Biotech Co., Ltd. (KR)
FEATURES
location/Qualifiers
source 1..762
/organism="Viscum album subsp. coloratum"
/db_xref="taxon:159976"
BASE COUNT 173 a 225 c 179 g 184 t 1 others
ORIGIN
Query Match 75.6%; Score 632.8; DB 6; Length 762;
Best Local Similarity 91.3%; Pred. No. 5.6e-179;
Matches 670; Conservative 1; Mismatches 63; Indels 0; Gaps 0;

Oy 1 CATCAGACAGCGGCGAGCATATTTCCGTTTCATTCACGCTTCCGAGATTAATGTCAC 60
|||
Db 28 CATCAACACAGCGGCGAGCATATTTCCGTTTCATTCACGCTTCCGAGATTAATGTCAC 87
|||
Oy 61 AGCGAAGCTTTTCCCATGAGATTAACATCTTCCGCTGAGTGTACAGATCCCGTCCGAT 120
|||
Db 88 AGCGAAGCTTTTCCCATGAGATTAACATCTTCCGCTGAGTGTACAGATCCCGTCCGAT 147
|||
Oy 121 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180
|||
Db 148 AGCGAAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 207
|||
Oy 181 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGAGCAATCCACTTTTGTG 240
|||
Db 208 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGAGCAATCCACTTTTGTG 267
|||
Oy 241 GCGGACGACACAGCGGCGGAAAGCATCTCTTCAACCGGACACACAGATCCTCTC 300
|||
Db 268 GCGGACGACACAGCGGCGGAAAGCATCTCTTCAACCGGACACACAGATCCTCTC 327
|||
Oy 301 CCATTCACGGAAGCTACACAGATCTGAGCGATACCGGTCATAGGAGACGATCCCT 360
|||
Db 328 CCATTCACGGAAGCTACACAGATCTGAGCGATACCGGTCATAGGAGACGATCCCT 387
|||
Oy 361 CTGGGTATAGAGAACTCATTCATCCGTCGCGCTTCTGTTATCCAGGCGGACGACAC 420
|||
Db 388 CTGGGTATAGAGAACTCATTCATCCGTCGCGCTTCTGTTATCCAGGCGGACGACAC 447
|||
Oy 421 GCGGACGACGCTGCTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 480
|||
Db 448 GCGGACGACGCTGCTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 507
|||
Oy 481 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGGCGGAGATCATTTCTCCGAC 540
|||
Db 508 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGGCGGAGATCATTTCTCCGAC 567
|||
Oy 541 ATGTACATCTGAGCTGAGACATGATTTGGGCGCCAAATCCAGCAATCCAGAGTCT 600
|||
Db 568 ACGTACATCTGAGCTGAGACATGATTTGGGCGCCAAATCCAGCAATCCAGAGTCT 627
|||
Oy 601 ACGATGAGGCTTTTAAATACCATTTCCGTTGGTAAATATCCACGCTTACCTTGTGACG 660
|||
Db 628 ACGATGAGGCTTTTAAATACCATTTCCGTTGGTAAATATCCACGCTTACCTTGTGACG 687
|||
Oy 661 TTGAGCAATGTCGACGCTGATCCGACGCTTACGATCTGTTGTATGAGGAGAC 720
|||
Db 688 TTGAGCAATGTCGACGCTGATCCGACGCTTACGATCTGTTGTATGAGGAGAC 747
|||
Oy 721 CGACCATCTTCTC 734
|||
Db 748 CGACCATCTTCTC 761
|||

RESULT 15
AX019439 762 bp DNA linear PAT 07-SEP-2000
LOCUS AX019439
DEFINITION Sequence 16 from Patent WO9940109.
ACCESSION AX019439
VERSION AX019439.1 GI:10043378
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequences.
REFERENCE 1 (bases 1 to 762)
AUTHORS Weilers,P., Stiefel,T., Voelter,W. and Morris,P.
TITLE Recombinant mistletoe lectins
JOURNAL Patent: WO 9940109-A 16 12-AUG-1999;
WEILERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH
(DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)
FEATURES
location/Qualifiers
source 1..762
/organism="synthetic construct"

PA (VISC-) VISCUM AG.
XX
XX
PI Kieff S:
XX
XX MPI; 2002-316737/36.
DR P-PSDB; ABB79450.
XX
XX New nucleic acid encoding preprotein of mistletoe lectin, useful as
PT diagnostic and therapeutic agents, also encodes polypeptide -
XX
PS Claim 1; Fig 1; 6pp; German.
XX
XX The invention relates to a nucleic acid molecule (AB156947) that encodes
CC a preprotein (ABB79450) which, after maturation, has the biological MLTII.
CC activity of the galactose-recognising mistletoe lectin (MLTII). The MLTII
CC activity of the galactose-recognising mistletoe lectin (MLTII). The MLTII
CC encoding nucleic acid molecule, primers specific to it or complements of
CC it, and encoded (oligomeric) polypeptides are useful as diagnostic and
CC therapeutic agents.
XX
XX Sequence 1656 BP; 387 A; 418 C; 449 G; 402 T; 0 other;

Query Match	94.78;	Score 792.8;	DB 24;	Length 1656;
Best Local Similarity	96.88;	Pred. No. 8,9e-249;		
Matches 809, Conservative	0;	Mismatches 27;	Indels 0;	Gaps 0;

QY	1	CATCGACGACGGGCGACGATATTTCCGGTTCAATACGCTTCTCCAGATATAGTCTCA	60
Db	127	CATCAACACAGGGCGAGCAATATTTCCGGTTCATACGCTTCTCCGAGTATATGTTCA	186
QY	61	AGCGGAGCTTTTCCATAGAGATACCACTCTTGCCGACATACAGTATCCCGTCTCGAT	120
Db	187	AGGGGAGGTTTTCATAGATATACCACTTTCCGTCATCTACAGATCCCGTCTCCGAT	246
QY	121	GGCGAAGATTTTGTCTGGTGGTGGAACTCACCAATCAGGGGGAGACTGCATACAGCGGCC	180
Db	247	GGCGAAGATTTTGTCTGGTGGAGCTCACCAATCAGGGGGAGACTGCATACGGCGCC	306
QY	181	ATGAGAGTACCAATCTGTAGCTGATGCTTACCAAGCAGGGGACCAATCTACTTTTG	240
Db	307	ATGAGAGTATCAATCTGTAGCTGATGCTTACCAAGCAGGGGACCAATCTACTTTTG	366
QY	241	CGCGAGCACACAGCGCGCGGAAAGCATCTCTCAACGGGACACAGATCTCTCTC	300
Db	367	CGCGAGCACATCTGACGGCGCGGAAAGGATCTCTCAACGGGACACAGATCTCTCTC	426
QY	301	CCATTCACCGAGATCACACAGATCTGGAGCGATACCGCGGTATAGGACACAGATCCT	360
Db	427	CCATTCACCGAGATCACACAGATCTGGAGGATACCGCGGTATAGGACACAGATCCT	486
QY	361	CTGGGTATAGGAACTCATTAATTCGTCTCGGCGCTGGTTTCAGGCGGACAGCAC	420
Db	487	CTGGGTATAGGAACTCATTAATTCGTCTCGGCGCTGGTTTCAGGCGGACAGCAC	546
QY	421	CGGGGCCAAGCTCCCTTATATCTCATTTCCAGATGATCTCCGAGCGCGGAAATTC	480
Db	547	CGGGGCCAAGCTCCCTTATATATCTCATTTCCAGATGATCTCCGAGCGCGGAAATTC	606
QY	481	AATCCCATCTTTTGGAGGCGTCCGCCAATACATTAACAGCGGGGAGATTAATTTCTCCGAC	540
Db	607	AATCCCATCTTTTGGAGGCGTCCGCCAAGATTAATACAGTGGGAGATTAATTTCTCCAGAC	666
QY	541	AATGTACATGCTCGAGCTGAGAGTACTGTTGGGGCAACAATCCAGCAAGTCCAGCAACTCT	600
Db	667	AATGTACATGCTCGAGCTGAGAGAGTGTGGGGCAACAATCCAGCAAGTCCAGCAACTCT	726
QY	601	ACGAGATGCGTTTAAATACCACTTTCCGTTGGGTATATCCACCGGTAATCTGTGTACG	660
Db	727	ACGAGATGCGTTTAAATACCACTTTCCGTTGGGTATATATCCACCGGTAATCTGTGTACG	786
QY	661	TTTAGCAATGTTCCGAGCTGATGCCAGGTATAGCAATCATGTTGTTGTATATGAGGAC	720
Db	787	TTTAGCAATGTTCCGAGCTGATGCCAGGTATAGCAATCATGTTGTTGTATGCGGGAC	846

QY	721	CGACGCAATCTTCCTCCGACGTCGCTATTGGCCGCTCTATATACGACGCGCTGTGGAAAT	780
Db	847	CGCGCAATCTTCCTCCGAATGCGCATTTGGCCCTGGTATACGACGCGTGTGGAAAT	906
QY	781	AGCGGCGCGCTCGAGACGCTTAACTCCGACGCGCTTCGCCAACCCACGCTGGCGATTCGT	836
Db	907	AGCGGCGCGCTCGAGACGCTTAACTCCGACGCGCTTCGCCAACCTACGCTGGCGATTCGT	962

RESULT 2
AAT91659
ID AAT91659 standard; DNA; 1923 BP.

AC AAT91659;

DT 18-DEC-1997 (first entry)

Prepro mistletoe lectin

XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss
KW

OS Viscum album.

Key	Location/Qualifiers
FF	1740

```

CBO      a
E1      /*tag=
FT      201

```

```

misc_readcare
FT
FT

```

misc feature	205..909
ET	/note- md gene
ET	

```

/*tag= c
/notes= "ML gene fragment f"

```

```

misc_feature      910.957

```

```

11 /note= "ML gene fragment e
12 FT
13
14 570 1746

```

```

*****
FI
FI
/*tag= e

```

misc feature	1747..1923
FTM	/noes
FTM	

```

1  /lag-
2  /note= "ML gene fragment 1"

```

XX

XX

XX
ED
XX

28 JUN 1955, 2000 0000
PE 2000 0000
XX

26-JUN-1993; 93EF 010000Z.
PR
XX

PA (MADU) MADATS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

AA
DR WPI; 1997-054678/06.

BR
E-PSDB, 10021;
XX

PT Nucleic acid encoding pre-pro form of
therapeutic or diagnostic use

XX
XX
300-1-1, Etc. AC: 3000: German.

XX
XX

...is an autotoxic agent that

It can be used in immunotoxins therapy.

CC acid fragments can be used in the synthesis of
CC AAT91659) comprises an A chain (AAT91660)
CC

XX
BP: 465 A; 488 C; 495 G; 47

03 18. score 695.2:

Best Local Similarity 90.78; Pred. No. 8.

Matches 120, conducted at 1000 ft.

OY 1 CATGACGACGCGGCGACGAAATATTCGGTTCATACGCGTTCGAGATTAATGCTCA 60
 DB 181 CATGAAACGACGCGGCGGAGAAATATTCGGTTCATACGCGTTCGAGATTAATGCTCA 240
 OY 61 AGCGGAAGCTTTTCCATGATGATACGCTCTTGGTCACTTACATCCCGTCTCGAT 120
 DB 241 AGCGGAAGCTTTTCCATGATGATACGCTCTTGGTCACTTACATCCCGTCTCGAT 180
 OY 121 GCGCAAGATTTGCTGCTGAGTACACTACCAATAGGCGGAGATCGATACGCGCC 300
 DB 301 GCGCAAGATTTGCTGCTGAGTACACTACCAAGGAGGAGATCGATACGCGCC 360
 OY 181 ATGAGCTTACCAATCTGTACGCTGAGTACCAAGGAGGAGATCGATACGCGCC 240
 DB 361 ATGAGCTTACCAATCTGTACGCTGAGTACCAAGGAGGAGATCGATACGCGCC 420
 OY 241 GCGGACGACGACGCGGCGGAGAAAGCATCTTACGCGGACGACGACGCGCGCTC 300
 DB 421 GCGGACGACGACGCGGCGGAGAAAGCATCTTACGCGGACGACGACGCGCGCTC 480
 OY 301 CCATTCACCGGAAGCTACAGATCTGAGGAGATACCGCGTATAGGAGCAATCCCT 360
 DB 481 CCATTCACCGGAAGCTACAGATCTGAGGAGATACCGCGTATAGGAGCAATCCCT 540
 OY 361 CTGGGTATAGAGAACTCATTCATCCGCTCGCGCTTCTGTTATCCAGGCGGACGAC 420
 DB 541 CTGGGTATAGAGAACTCATTCATCCGCTCGCGCTTCTGTTATCCAGGCGGACGAC 500
 OY 421 GCGGCGCAAGCTCTGCTTATATATCTCATTCAGATGATCTCGGAGCGCGAGATT 480
 DB 601 GGTACCAAGCTCTGCTTATATATCTCATTCAGATGATCTCGGAGCGCGAGATT 560
 OY 481 AATCCATCTTTTGGAGGCTCGCCATATACATTAACGCGGAGCTATCTTCTCCGAC 540
 DB 661 AATCCATCTTTTGGAGGCTCGCCATATACATTAACGCGGAGCTATCTTCTCCGAC 620
 OY 541 ATGTACATCTGAGCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG 600
 DB 721 ATGTACATCTGAGCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG 780
 OY 601 AGCGATGCTTTTAAATTAATCCATTTGCTGATATCCACCGCTTACTTCTGAGC 660
 DB 781 AGCGATGCTTTTAAATTAATCCATTTGCTGATATCCACCGCTTACTTCTGAGC 840
 OY 661 TTGAGCAATCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 720
 DB 841 TTGAGCAATCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 900
 OY 721 GAGCATCTCTCCGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 780
 DB 901 GAGCATCTCTCCGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 960
 OY 781 AGCGGCGCGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 836
 DB 954 AGCGGCGCGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 1001
 OY 954 AGCGGCGCGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 1001

FH Key Location/Qualifiers
 FT CDS 55..1749
 FT /tag- a
 FT /product- "mistletoe lectin"
 EP884388-21.
 16-DEC-1998.
 26-JUN-1995; 95EP-0109949.
 26-JUN-1995; 95EP-0109949.
 26-JUN-1995; 98EP-0105660.
 (MADU) MADUS KOEHL AG.
 Baur A, Eck J, Lentzen H, Zinke H;
 MPI; 1999-02582/03.
 P-FSDB; AAW90127.
 New transgenic plant expressing mistletoe lectin - useful for
 producing recombinant lectin in e.g. cancer diagnosis and therapy
 Claim 1a; Fig 4c; 30pp; German.
 This invention describes a novel transgenic plant transformed with a
 vector capable of encoding a mistletoe (Viscum album) lectin
 preprotein or a biologically active fragment. The specific lectin
 also describes a polypeptide produced by a plant where the polypeptide
 exhibits at least one enzymatic modification other than the glycosylation
 that occurs in Viscum album or the polypeptide is a fusion protein, a
 mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 polypeptide or the polypeptide dimer. The plants are used for large-scale
 production of mistletoe lectin for diagnostic or therapeutic purposes
 (e.g. in cancer therapy). This sequence encodes the mistletoe lectin
 used in the method of the invention.
 Sequence 1923 BP; 465 A; 488 C; 495 G; 475 T; 0 other:
 Query Match 83.1%; Score 695.2; DB 20; Length 1923;
 Best Local Similarity 90.7%; Pred. No. 8.5e-217;
 Matches 758; Conservative 0; Mismatches 63; Indels 15; Gaps 1;
 OY 1 CATGACGACGCGGCGACGAAATATTCGGTTCATACGCGTTCGAGATTAATGCTCA 60
 DB 181 CATGAAACGACGCGGCGGAGAAATATTCGGTTCATACGCGTTCGAGATTAATGCTCA 240
 OY 61 AGCGGAAGCTTTTCCATGATGATACGCTCTTGGTCACTTACATCCCGTCTCGAT 120
 DB 241 AGCGGAAGCTTTTCCATGATGATACGCTCTTGGTCACTTACATCCCGTCTCGAT 180
 OY 121 GCGCAAGATTTGCTGCTGAGTACACTACCAATAGGCGGAGATCGATACGCGCC 300
 DB 301 GCGCAAGATTTGCTGCTGAGTACACTACCAAGGAGGAGATCGATACGCGCC 360
 OY 181 ATGAGCTTACCAATCTGTACGCTGAGTACCAAGGAGGAGATCGATACGCGCC 240
 DB 361 ATGAGCTTACCAATCTGTACGCTGAGTACCAAGGAGGAGATCGATACGCGCC 420
 OY 241 GCGGACGACGACGCGGCGGAGAAAGCATCTTACGCGGACGACGACGCGCGCTC 300
 DB 421 GCGGACGACGACGCGGCGGAGAAAGCATCTTACGCGGACGACGACGCGCGCTC 480
 OY 301 CCATTCACCGGAAGCTACAGATCTGAGGAGATACCGCGTATAGGAGCAATCCCT 360
 DB 481 CCATTCACCGGAAGCTACAGATCTGAGGAGATACCGCGTATAGGAGCAATCCCT 540
 OY 361 CTGGGTATAGAGAACTCATTCATCCGCTCGCGCTTCTGTTATCCAGGCGGACGAC 420
 DB 541 CTGGGTATAGAGAACTCATTCATCCGCTCGCGCTTCTGTTATCCAGGCGGACGAC 500
 OY 421 GCGGCGCAAGCTCTGCTTATATATCTCATTCAGATGATCTCGGAGCGCGAGATT 480

```

Db 601 CGTACCAAGCTCGTTCGATTTAAATCCATTCAGATGATTCGAGAGCCGCCAGATG 660
Qy 481 AATCCCATCTTTGGAGGCTCCCAATACATTAACAGCGAGAGCATTTCTCCCGAC 540
Db 661 AATCCCATCTTTGGAGGCTCCCAATACATTAACAGCGAGAGCATTTCTCCCGAC 720
Qy 541 ATGTACATGCTCGAGTGAAGTATGAGGCGCCAAATCCAGCAGATCCAGACATCT 600
Db 721 GTGTACATGCTCGAGTGAAGTATGAGGCGCCAAATCCAGCAGATCCAGACATCT 780
Qy 601 ACGATGAGCGTTTAAATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGAC 660
Db 781 ACGATGAGCGTTTAAATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGAC 840
Qy 661 TTGAGCAATGTTGCGAGCGATGATCGCCAGCTTACGATCAATGTTGTATGAGGAC 720
Db 841 TTGAGCAATGTTGCGAGCGATGATCGCCAGCTTACGATCAATGTTGTATGAGGAC 900
Qy 721 CGACATCTTCCTCCGACGTGCGCTAATGCGCGCTGCTGCTATACAGACCCTCTTGAAAT 780
Db 901 CGACATCTTCCTCCGACGTGCGCTAATGCGCGCTGCTGCTATACAGACCCTCTTGAT 953
Qy 781 AGCGGCGCGTCCGACGATGTTACCTGACATGCTCTCCGAAACCCAGCGTGCAGTCT 836
Db 954 -----AGCGATGATGTTACCTGACATGCTCTCCGAAACCCAGCGTGCAGTCT 1001

RESULT 4
AAZ09103 standard: DNA; 1596 BP.
AAZ09103;
18-OCT-1999 (first entry)
Mistletoe lectin I DNA fragment.
Mistletoe lectin: antitumor; immunostimulant; A-chain; MIA; immunity;
ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response;
lymphokine-producing macrophage; uncontrolled cell growth; treatment;
cancer; cytotoxicity; antigen; isoform; lectin I; ds.
Vascum album.
De19804210-A1.
12-AUG-1999.
03-FEB-1998: 98DE-1004210.
03-FEB-1998: 98DE-1004210.
03-FEB-1998: 98DE-1004210.
(BIOS-) BIOSYN ARZNEIMITTEL GMBH.
Morris P, Stiefel T, Voelter W, Walters P;
WPI; 1999-44535/38.
P-PSDB; AAY25979.
Preparation of mistletoe lectins in heterologous systems,
particularly for use as anticancer agents and immunostimulants
Claim 14; Fig 1A; 78pp; German.
This invention describes a novel mistletoe lectin (I) and its fragments
which have antitumor and immunostimulatory activity. The A-chain (MIA)
of the mistletoe lectin binds to, and inactivates the 28S subunit of
ribosomes. Non-cytotoxic forms of (I) activate T-cell and
lymphokine-producing macrophages, so stimulate immunity. (I) and its
fragments are used to treat uncontrolled cell growth (particularly
cancers) and if they lack cytotoxicity to increase the strength of the
immune response, particularly to a co-administered antigen

```

```

Cc (tumour-associated, bacterial or viral). The method allows production of
Cc mistletoe lectin, and its individual chains, in many different isoforms
Cc and on a large scale, at any time of the year. Recombinant products are
Cc free from toxins present in natural mistletoe extracts. This sequence
Cc encodes a fragment of the mistletoe lectin I protein.
Qy 50
Sequence 1596 BP; 387 A; 418 C; 419 G; 372 T; 0 other;
Query Match 82.7%; Score 692; DB 20; Length 1596;
Best Local Similarity 90.4%; Pred. No. 8.7e-216;
Matches 756; Conservative 0; Mismatches 65; Indels 15; Gaps 1;
Qy 1 CATCAGACGAGGCGGCGAGCAATATTTCCGTTCCAGCTTCTCCGAGATATGATCA 60
Db 28 CATCAACACGAGGCGGCGAGCAATATTTCCGTTCCAGCTTCTCCGAGATATGATCA 87
Qy 61 AGCGAAGCTTTTCCAAATGAGATACCACTTTCGTCAGTACAGATCCGCTCGAT 120
Db 88 AGCGAAGCTTTTCCAAATGAGATACCACTTTCGTCAGTACAGATCCGCTCGAT 147
Qy 121 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180
Db 148 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 207
Qy 181 ATGACATTTACCAATCTGTACGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 240
Db 208 ATGACATTTACCAATCTGTACGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 267
Qy 241 GCGGACGACACGAGCGCGCGGAAAGCATCTCTCCAGCGGACACAGATCTCTCTC 300
Db 268 GCGGACGACACGAGCGCGCGGAAAGCATCTCTCCAGCGGACACAGATCTCTCTC 327
Qy 301 CCATTCACCGCAACCTTACAGATCTGAGAGCATAGCGGCTATAGGAGCAAGATCT 360
Db 328 CCATTCACCGCAACCTTACAGATCTGAGAGCATAGCGGCTATAGGAGCAAGATCT 387
Qy 361 CTGCGATGAGAGCAATCATATCCCTGCGCGCTTCTGTTACAGAGCGGACAGAC 420
Db 388 CTGCGATGAGAGCAATCATATCCCTGCGCGCTTCTGTTACAGAGCGGACAGAC 447
Qy 421 CGCGCGCAAGGCTGCTCCCTTAAATCTCATGATGATCTCGAGCGCGGAGATTC 480
Db 448 CGTACCAAGCTCTGATTTAACTTCATGATGATCTCGAGCGCGGAGATTC 507
Qy 481 AATCCCATCTTTGGAGGCTCCCAATACATTAACAGCGAGAGCATTTCTCCCGAC 540
Db 508 AATCCCATCTTTGGAGGCTCCCAATACATTAACAGCGAGAGCATTTCTCCCGAC 567
Qy 541 ATGATACGCTCGAGTGAAGTATGAGGCGCCAAATCCAGCAGATCCAGATCT 600
Db 568 GTGTACATGCTCGAGTGAAGTATGAGGCGCCAAATCCAGCAGATCCAGATCT 627
Qy 601 ACGATGAGCGTTTAAATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGAC 660
Db 628 ACGATGAGCGTTTAAATTAACCCATTTGCGTTGGGTATATCCACCGGTAACTTCGTGAC 687
Qy 661 TTGAGCAATGTTGCGAGCGATGATCGCCAGCTTACGATCAATGTTGTATGAGGAC 720
Db 688 TTGAGCAATGTTGCGAGCGATGATCGCCAGCTTACGATCAATGTTGTATGAGGAC 747
Qy 721 CGACATCTTCCTCCGACGTGCGCTAATGCGCGCTGCTGCTATACAGACCCTCTTGAAAT 780
Db 748 CGACATCTTCCTCCGACGTGCGCTAATGCGCGCTGCTGCTATACAGACCCTCTTGAT 800
Qy 781 AGCGGCGCGTCCGACGATGTTACCTGACATGCTCTCCGAAACCCAGCGTGCAGTCT 836
Db 801 -----AGCGATGATGTTACCTGACATGCTCTCCGAAACCCAGCGTGCAGTCT 1001

RESULT 5
AAZ09100 standard: DNA; 1598

```

AC AA209100;
XX
XX 18-OCT-1999 (first entry)
XX
XX Mistletoe lectin DNA consensus sequence.
DE
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KM riboxyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KM cancer; cytotoxicity; antigen; Isoform; ss.
XX
XX Viscum album.
OS
XX DE19804210-A1.
PN
XX 12-ANG-1999.
PD
XX 03-FEB-1998; 98DE-1004210.
PF
XX 03-FEB-1998; 98DE-1004210.
PR
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PA
XX Morris P, Stiefel T, Voelter W, Welters P;
PI
XX MPI; 1999-445335/38.
DR
XX
XX Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
XX
XX Claim 11; Page 32-33; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of mistletoe lectin DNA described in the
CC specification.
XX
XX Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;
SQ
Query Match 81.7%; Score 683.8; DB 20; Length 1598;
Best Local Similarity 87.2%; Pred. No. 4.2e-213;
Matches 730; Conservative 37; Mismatches 54; Indels 16; Gaps 2;
OY 1 CATCAGACGAGCGGCGACGAATATTCGCGTTCATCAGCGCTTCGAGATATGTCCTCA 60
DB 28 CATCAACACCGACGGCGAKGAACTCTCGGTTCATCAGCGCTTCGAGATATGTCCTCA 87
OY 61 AGCGGAAGCTTTTCCATATGATACCACTTCGCGTACGTCAGTACCCCGCTCGGAT 120
DB 88 AGCGGAAGCTTTTCCATATGATACCACTTCGCGTACGTCAGTACCCCGCTCGGAT 147
OY 121 GCGCAAGATTTTGTGTGGAAGTACCAATCAGGAGGAGATCGATCAGGCGGCC 180
DB 148 GCGCAAGATTTTGTGTGGAAGTACCAATCAGGAGGAGATCGATCAGGCGGCC 207
OY 181 ATCGACGTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCTTTTG 240
DB 208 ATCGACGTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCTTTTG 267
OY 241 GCGCAGCAGCAGCGCGCGGAAAGCATCTTCAACCGGACACACCAAGA-TCTCTCTT 299
DB 268 GCGCAGCAGCAGCGCGCGGAAAGCATCTTCAACCGGACACACCAAGA-TCTCTCTT 327

OY 300 CCCATTACCGGAGAGCTACAGATCTGAGCATACCGCGGTATAGGACAGATCCC 359
DB 328 CCCATTACCGGAGAGCTACAGATCTGAGCATACCGCGGTATAGGACAGATCCC 387
OY 360 TCTGGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTATACGAGCGCAGAC 419
DB 388 TCTCGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTATACGAGCGCAGAC 447
OY 420 CCGGCGCCAGCGTCTGCTCCCTATATACCTCATTCAGATGATGTCGCGGCGGAGATT 479
DB 448 GCGTTCYAAAGCTCGTGTGATTTATCTCTCATTCAGATGATGTCGCGGCGGAGATT 507
OY 480 CATCCCATCTTTTGGAGGCGCTCGCAATACATTAACAGCGGAGATCTTCTCCGA 539
DB 508 CATCCCATCTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
OY 540 CATGTACATGCTCGAGCTGAGACTAGTTGGGCGCCAAATCCAGCAGATCCAGAGTC 599
DB 568 CRTGTACATGCTCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
OY 600 TACGAGTGGCGCTTTTATTAACCCATTCGCTGGGTATATCCACCGGTACTCTGAGAC 659
DB 628 AACCGATGCGCTTTTATTAACCCATTCGCTGGGTATATCCACCGGTACTCTGAGAC 687
OY 660 GTTGAGCAATGTTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTATGAGGGA 719
DB 688 GTTGAGCAATGTTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTATGAGGGA 747
OY 720 CCGACCATCTTCTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTATGAGGGA 779
DB 748 GCGGCGCATCTTCTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTATGAGGGA 801
OY 780 TAGCGCGCGCTTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTATGAGGGA 836
DB 802 -----AGCGGATGATGTTACTGCAAGTGTCTGGAAGCTTACGATGAGGATTTGT 849
RESULT 6
ID AAC85472 standard; cDNA; 762 BP.
AAC85472;
XX
XX AAC85472;
AC
XX 16-MAY-2001 (first entry)
DT
XX
XX A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #1.
DE
XX
XX Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KM KML; tumour; KM-110; KML-C; KMLBP; KML-ITU; KML-ILU;
KM heparin binding protein; ds.
XX
XX Viscum album coloratum.
OS
XX EPI074560-A2.
PN
XX
XX 07-FEB-2001.
PD
XX
XX 27-JUL-2000; 2000EP-0402158.
PF
XX
XX 27-JUL-1999; 99KR-0030638.
PR
XX
XX (MIST-) MISTLE BIOTECH CO LTD.
PA
XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;
PI
XX
XX MPI: 2001-171044/18.
DR
XX P-PSDB; AAB47090.
PT
XX Novel lectin proteins isolated from Korean mistletoe, useful for
PT enhancing immunity and effectuating anti-tumoral activity -
PS
Claim 1; Page 25; 62pp; English.

XX The sequences given in AAC854/72-74 encode isoforms of an A-chain gene
CC involved in biosynthesis of lectins isolated from Korean mistletoe.
CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
CC for treating tumours. The KML's are isolated from a protein fraction
CC derived from the leaves, stems and fruits of Korean mistletoe, which
CC is designated KM-110. One of the isolates, KML-C was shown to be
CC effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.

Sequence 762 BP; 172 A; 220 C; 186 G; 184 T; 0 other;

Query Match	79.98;	Score 668.4;	DB 22;	Length 762;
Best Local Similarity	94.48;	Pred. No. 3.3e-208;		
Matches 693; Conservative	0;	Mismatches 41;	Indels 0;	Gaps 0;

QY	1	CATGAGCGACGGGCGACGAATTTTCGGTTCATACAGCTTCCTCCGAGATTATGTCTCA	60
Db	28	CATTAACCAACAGGGCGACGAATTTTTCGGTTCATCAAGCTTCCTCCGAGATTATGTCTCA	87
QY	61	AGCGGAAGCTTTTTCGAATGAGATACCACTTCCTCGTCACTACAGATCCCGCTCTGGAT	120
Db	88	AGCGGAAGCTTTTTCGAATGAGATACCACTTCCTCGTCACTACAGATCCCGCTCTGGAT	147
QY	121	GCGGAAGATTGTGTGGTGGGAACCTACCAATCAGGGGGGAAACTCGATCAGCGCGCC	180
Db	148	GCGGAAGATTGTGTGGTGGGAACCTACCAATCAGGGGGGAAACTCGATCAGCGCGCC	207
QY	181	ATCGAGCTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCACTTTTG	240
Db	208	ATCGAGCTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCACTTTTG	267
QY	241	CGGAGCGACCAAGCGGGCGGGAAGGCACTCTTACCGGGCACACAGATCTCTTC	300
Db	268	CGGAGCGACCAAGCGGGCGGGAAGGCACTCTTACCGGGCACACAGATCTCTCTC	327
QY	301	CCATTACCGGAAGCTACACAGATCTGGAGCGGATGCGGTCATAGGAGCACAGATCCCT	360
Db	328	CCATTACCGGAAGCTACACAGATCTGGAGCGGATGCGGTCATAGGAGCACAGATCCCT	387
QY	361	CTGGGTATGAGGAGCACTTCAATCCGTCGCGGCGTTGTATATCAGGCGGCAACAC	420
Db	388	CTGGGTATGAGGAGCACTTCAATCCGTCGCGGCGTTGTATATCAGGCGGCAACAC	447
QY	421	CGGGCCCAAGCTGTTCCCTTATATATCTTCATTCAGATGATCTCCAGGCGCGAGATTC	480
Db	448	CGTGCCCAAGCTGTTCCCTTATATATCTTCATTCAGATGATCTCCAGGCGCGAGATTC	507
QY	481	AATGCCATCTTTTGGAGGGGTCCGCAATACATTAAACGCGGGAGATCTTCTCCGAC	540
Db	508	AATGCCATCTTTTGGAGGGGTCCGCAATACATTAAAGAGTGGGGGGTCAATTTCTGCCAGAC	567
QY	541	ATGTACATGCTCAGCTGAGACTAGTTGGGGCAACAATCCAGCGAAGTCAGCAGTCT	600
Db	568	ACGTACATTTCCAGCTGGAGAGGAGTGGGGCAACAATCCAGCGAAGTTCAGCAGTCTG	627
QY	601	ACGGATGGCGTTTATATACCATTTCGTTGGGTATATCCACCGGTAACTTCGTGACG	660
Db	628	ACGGATGGCGTTTATATACCATTTCGTTGGGTATATCCACCGGTAACTTCGTGAGAC	687
QY	661	TTGAGCAATTTTCGACAGCATCGCCAGCTTACGATATGTTTGTATATAGGAC	720
Db	688	TTGAGCAATTTTCGACAGCATCGCCAGCTTACGATATGTTTGTATATCCAGGAC	747
QY	721	CGACCATCTTCCTC	734
Db	748	CGGCAATCTTCCTC	761

RESULT 7	
AAT91660	
ID	AAT91660 standard; DNA; 774 BP.
XX	
AC	AAT91660;

XX	18-DEC-1997	(first entry)
DT		
XX	Prepro mistletoe lectin A chain.	
DE		
XX	Mistletoe: lectin; cytotoxic; A chain; B chain; dimer; ss.	
KW		
XX	Viscum album.	
OS		
Key	Location/Qualifiers	
FH	4..765	
FT		
FT	/**tag= a	
XX		
EN	EP751221-A1.	
XX		
PD	02-JAN-1997.	
XX		
PE	26-JUN-1995; 95EP-0109949.	
XX		
PR	26-JUN-1995; 95EP-0109949.	
XX		
PA	(MADU) MAD AUS KOELN AG.	
XX		
PI	Baur A, Eck J, Lentzen H, Zinke H;	
DR	WPI; 1997-054678/06.	
DR	P-PSDB; AAW10022.	
XX		
PT	Nucleic acid encoding prepro form of mistletoe lectin - for	
PT	therapeutic or diagnostic use	
XX		
PS	Claim 2; Fig 4A; 30pp; German.	
XX		
CC	Mistletoe lectin is a cytotoxic agent that has been used for tumour	
CC	therapy. It can be used in immunotoxins and medicaments. Nucleic	
CC	acid fragments can be used in diagnostic methods. Mistletoe lectin (
CC	AA191659) comprises an A chain (AA191660) and a B chain (AA191661).	
XX		
SO	Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other;	

Query Match	77.1%;	Score 645.2;	DB 18;	Length 774;
Best Local Similarity	92.7%;	Pred. No. 1.3e-20;		
Matches 677;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;
QY 1	CATCAGACGAGGGCGACGAAATATTTCCGGTTATCATCAGGCTTCTCCGAGATATTGTCGA	60		
Db 34	CACCAACACACCGGTGAAGAATATTTCCGGTTCATCAGGCTTCTCCGAGATATTGTCGA	93		
QY 61	AGCGGAGCTTTTCCATGAGATACCACTCTTGCGTCAGTTCAGATATCCCGCTCGGAT	120		
Db 94	AGCGGAAGCTTTTCCATGAGATACCACTCTTGCGTCAGTTCAGATATCCCGCTTCGAT	153		
QY 121	GCAGGAAGATTTGTTGTGTGGAACATCACCAATCAGGGGGAGACATCGATACAGCGCGC	180		
Db 154	GCAGGAAGATTTGTTGTGTGGAACATCACCAATCAGGGGGAGACATCGATACAGCGCGC	213		
QY 181	ATCGAGCTTACCAATCTGTACGTGTGTGCTTACCAAGCAGGCGACCAATCTACTTTTG	240		
Db 214	ATCGAGCTTACCAATCTGTACGTGTGTGCTTACCAAGCAGGCGACCAATCTACTTTTG	273		
QY 241	CGGAGCAGCAGACGAGCGGCGGAAAGCAATCTTACCGGGACACAGATCCGTCGC	300		
Db 274	CGGAGCAGCAGACGAGCGGCGGAAAGCAATCTTACCGGGACACAGATCCGTCGC	333		
QY 301	CCATTCCCGGAAGCTACACAGATCTGGAGCGATACGCGGCTATAGGGACAGATCCCT	360		
Db 334	CCATTCCCGGAAGCTACCTGTATGGAAGCATATAGCGCGGACATAGGGACAGATCCCT	393		
QY 361	CTGGGATTAAGAGAACTATTCAATCCGCTCTGGCGCTTCGTATTCAGGCGGACAGACC	420		
Db 394	CTGGGATTAAGAGAACTATTCAATCCGCTCTGGCGCTTCGTATTCAGGCGGACAGACC	453		
QY 421	CGGGCCCAACCTGTTTCCCTTATATATCCATTCAGATGATCTCCGAGGCGGAGATTTC	480		

PT Preparation of mistletoe lectins in heterologous systems,
 particularly for use as anticancer agents and immunostimulants
 PS Claim 12; Page 34; 78pp; German.
 XX

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumor and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
 CC described in the specification.
 XX

XX Sequence 763 BP; 163 A; 215 C; 175 G; 170 T; 40 other;

Query Match 76.68; Score 641; DB 20; Length 763;
 Best Local Similarity 89.18; Pred. No. 3,1e-199;
 Matches 655; Conservative 37; Mismatches 42; Indels 1; Gaps 1;

QY 1 CATCAGACGACGGGCGACATATTCGGTTCATCAGCGTCTCCGAGATTATGCTCA 60
 DB CATCAACACGCGGCGGCGGAGATCTTCGGTTCATCAGCGTCTCCGAGATTATGCTCA 87
 QY 61 AGCGGAAGCTTTTCCAAATGAGATACCACTTCGTCAGTCTACGATCCCGCTCTCGAT 120
 DB AGCGGAAGCTTTTCCAAATGAGATACCACTTCGTCAGTCTACGATCCCGCTCTCGAT 147
 QY 88 AGCGGAAGCTTTTCCAAATGAGATACCACTTCGTCAGTCTACGATCCCGCTCTCGAT 147
 DB AGCGGAAGCTTTTCCAAATGAGATACCACTTCGTCAGTCTACGATCCCGCTCTCGAT 147
 QY 121 GCGCAAGATTTGTTGGTGGAACTCACCACATCAGGGGGGAGACTGATCAGCGCGCC 180
 DB GCGCAAGATTTGTTGGTGGAACTCACCACATCAGGGGGGAGACTGATCAGCGCGCC 207
 QY 148 GCGCAAGATTTGTTGGTGGAACTCACCACATCAGGGGGGAGACTGATCAGCGCGCC 207
 DB GCGCAAGATTTGTTGGTGGAACTCACCACATCAGGGGGGAGACTGATCAGCGCGCC 207
 QY 181 ATGACGTTACCAATCTGATCGTGGTGGTTCACAGACGAGCGCAATCTCTTTTG 240
 DB ATGACGTTACCAATCTGATCGTGGTGGTTCACAGACGAGCGCAATCTCTTTTG 267
 QY 208 ATGACGTTACCAATCTGATCGTGGTGGTTCACAGACGAGCGCAATCTCTTTTG 267
 DB ATGACGTTACCAATCTGATCGTGGTGGTTCACAGACGAGCGCAATCTCTTTTG 267
 QY 241 GCGGACGACACGAGCGCGGAAAGGACTCTTCACGCGGACACAGCA-TCTCTCTG 299
 DB GCGGACGACACGAGCGCGGAAAGGACTCTTCACGCGGACACAGCA-TCTCTCTG 327
 QY 268 GCGGACGACACGAGCGCGGAAAGGACTCTTCACGCGGACACAGCA-TCTCTCTG 327
 DB GCGGACGACACGAGCGCGGAAAGGACTCTTCACGCGGACACAGCA-TCTCTCTG 327
 QY 300 CCATTCACGGAAGCTACACATCTGAGAGCATACGCCGTCATAGGAGACAGATCC 359
 DB CCATTCACGGAAGCTACACATCTGAGAGCATACGCCGTCATAGGAGACAGATCC 387
 QY 328 CCATTCACGGAAGCTACACATCTGAGAGCATACGCCGTCATAGGAGACAGATCC 387
 DB CCATTCACGGAAGCTACACATCTGAGAGCATACGCCGTCATAGGAGACAGATCC 387
 QY 360 TCTGGTATAGAGAACTCATTCACGCTCGGCGCTTCGTTATCCAGGGCGAGCAC 419
 DB TCTGGTATAGAGAACTCATTCACGCTCGGCGCTTCGTTATCCAGGGCGAGCAC 447
 QY 388 TCTGGTATAGAGAACTCATTCACGCTCGGCGCTTCGTTATCCAGGGCGAGCAC 447
 DB TCTGGTATAGAGAACTCATTCACGCTCGGCGCTTCGTTATCCAGGGCGAGCAC 447
 QY 420 CCGGCGCAAGCTGTTCCCTTAATCTCATATTCAGATATCCGAGCGCCGAGATT 479
 DB CCGGCGCAAGCTGTTCCCTTAATCTCATATTCAGATATCCGAGCGCCGAGATT 507
 QY 448 GCGTTCACAGCTGCTGCTGATTTTATCTCATTCAGATATCCGAGCGCCGAGATT 507
 DB GCGTTCACAGCTGCTGCTGATTTTATCTCATTCAGATATCCGAGCGCCGAGATT 507
 QY 480 CAATCCCATTTTGGAGGGCTCGGCAATATTAACATACGCGGGGATCTTTTCCCA 539
 DB CAATCCCATTTTGGAGGGCTCGGCAATATTAACATACGCGGGGATCTTTTCCCA 567
 QY 508 CAATCCCATTTTGGAGGGCTCGGCAATATTAACATACGCGGGGATCTTTTCCCA 567
 DB CAATCCCATTTTGGAGGGCTCGGCAATATTAACATACGCGGGGATCTTTTCCCA 567
 QY 540 CATGATCATGCTGAGAGCTGAGATGTTGGGCGCAACATCCAGCAGATCCAGAGTC 599
 DB CATGATCATGCTGAGAGCTGAGATGTTGGGCGCAACATCCAGCAGATCCAGAGTC 627
 QY 568 CATGATCATGCTGAGAGCTGAGATGTTGGGCGCAACATCCAGCAGATCCAGAGTC 627
 DB CATGATCATGCTGAGAGCTGAGATGTTGGGCGCAACATCCAGCAGATCCAGAGTC 627
 QY 600 TAGGATGGGCTTTTAATTAACCATTTGGGTTATATCCAGCGGTAATCTCTGAC 659
 DB TAGGATGGGCTTTTAATTAACCATTTGGGTTATATCCAGCGGTAATCTCTGAC 687
 QY 628 AACCGATGGGCTTTTAATTAACCATTTGGGTTATATCCAGCGGTAATCTCTGAC 687
 DB AACCGATGGGCTTTTAATTAACCATTTGGGTTATATCCAGCGGTAATCTCTGAC 687
 QY 660 GTTGACCAATGTTGCGAGTGATCCGCAAGTATAGCATCATGTTTGTATGATGAGGA 719
 DB GTTGACCAATGTTGCGAGTGATCCGCAAGTATAGCATCATGTTTGTATGATGAGGA 747
 QY 688 GTTGACCAATGTTGCGAGTGATCCGCAAGTATAGCATCATGTTTGTATGATGAGGA 747
 DB GTTGACCAATGTTGCGAGTGATCCGCAAGTATAGCATCATGTTTGTATGATGAGGA 747

QY 720 CCAGCATCTCTCTC 734
 DB 748 GCGGCGATCTCTCTC 762

RESULT 11
 AAV51341
 ID AAV51341 standard; DNA; 762 BP.
 AC AAV51341;
 XX 23-OCT-1998 (first entry)
 DE Mistletoe rMLA DNA.
 XX
 KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
 XX
 OS Viscum album.
 XX
 XX Key Location/Qualifiers
 FH 4..759
 FT CDS
 FT /*tag= a
 FT /product= "Lectin A-chain"
 FT /note= "Partial sequence"
 XX
 PN W09829540-A2.
 XX
 PD 09-JUL-1998.
 XX
 PF 02-JAN-1998; 98WO-EP00009.
 XX
 PR 02-JAN-1997; 97EP-010012.
 XX
 PA (BRAI-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
 XX
 PI Eck J, Schmidt A, Zinke H;
 XX
 DR WPI; 1998-388122/33.
 XX
 DR P-PSDB; AAM64659.
 XX
 PT Nucleic acid encoding fusion protein containing mistletoe lectin A
 PT chain - useful for treatment of proliferative and autoimmune
 PT diseases, allergies and tumours
 XX
 PS Disclosure; Fig 11a; 115pp; German.
 XX

This sequence encodes a lectin A-chain, rMLA, isolated from mistletoe.
 This sequence can be used in the construction of a fusion protein which
 comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC e.g. by injection or topically but especially by intravenous injection,
 CC at 1 ng to 500 mg g/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.
 XX

Sequence 762 BP; 169 A; 230 C; 183 G; 180 T; 0 other;

Db 568 ACGTACATTCCTCAGCTGGAGACGAGTGGGGGCAACATCCAGCAAGTCCACAGACTCG 627
 Qy 601 ACGATGGCGTTTAAATAAACCCATTGGGTATATCCACCGTAACTTCGTGACG 660
 Db 638 ACGGATGGCGTTTAAATAAACCCATTGGGTATATCCACCGTAACTTCGTGACG 687
 Qy 661 TTGAGCAATGTTCCGAGCATGCGCAGCTTACGATCATGTTGTATGTATGAGGAC 720
 Db 668 TTGAGCAATGTTCCGAGCATGCGCAGCTTACGATCATGTTGTATGTATGAGGAC 747
 Qy 721 CGACCATCTTCCTC 734
 Db 748 CGGCATCTTCCTC 761

RESULT 13

AAZ09104
ID AAZ09104 standard; DNA; 762 BP.

AAZ09104;

DE 18-OCT-1999 (first entry)

DE Mistletoe lectin A1 DNA fragment.

KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A1; ds.
 OS Viscum album.

XX DE19804210-A1.

PN 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.

XX 03-FEB-1998; 98DE-1004210.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX PI Morris P, Stiefel T, Voelter W, Welters P;

XX WPI; 1999-445335/38.

XX DR P-PDB; AAY25980.

XX PT Preparation of mistletoe lectins in heterologous systems,
 particularly for use as anticancer agents and immunostimulants

XX PS Claim 15; Fig 2A; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (II) and its
 CC fragments are used to treat uncontrolled cell growth (particularly of
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC encodes a fragment of the mistletoe lectin A1 protein.

SO Sequence 762 BP; 171 A; 230 C; 183 G; 178 T; 0 other;

Query Match 75.5%; Score 631.6; DB 20; Length 762;
 Best Local Similarity 91.3%; Pred. No. 3.7e-196;
 Matches 670; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 1 CATCAGACGACGGGCGACGAAATATTCGGTTATACGCTTCCGAGATTATGCTCA 60
 Db 28 CATCAAAACCGGCGGCGAGAACTCCGGTTATACGCTTCCGAGATTATGCTCA 87
 Qy 61 AGCGAAGCTTTTCCATGATGATACCACTTCGTCAGTACGATACGATCCCGTCGGAT 120
 Db 88 AGCGAAGCTTTTCCATGATGATACCACTTCGTCAGTACGATACGATCCCGTCGGAT 147
 Qy 121 GCGCAAAAGATTTGTGGTGTGAACTACCAATCAGGGGAGACTGATCAAGCGCGC 180
 Db 148 GCGCAAAAGATTTGTGGTGTGAACTACCAATCAGGGGAGACTGATCAAGCGCGC 207
 Qy 181 ATGAGGTTACCAATCTGATGCTGCTTACCAAGCAGGAGACCAATCTATTG 240
 Db 208 ATGAGGTTACCAATCTGATGCTGCTTACCAAGCAGGAGACCAATCTATTG 267
 Qy 241 CGGAGACGACGAGCGGCGGAAAGCAATCTTCACGCGGACGACGATCTCTC 300
 Db 268 CGGAGACGACGAGCGGCGGAAAGCAATCTTCACGCGGACGACGATCTCTC 327
 Qy 301 CCATTACCGGAGAGCTACACAGATCTGAGCGATACGCGGTCATAGGAGACGATCCCT 360
 Db 328 CCATTACCGGAGAGCTACCGCTGATCTGAGCGATACGCGGATAGGAGACGATCCCT 387
 Qy 361 CTGGGATAGAGAACTCATTCATCTGCTGCGCTTCGTTATCCAGGCGGACGACC 420
 Db 388 CTGGGATAGAGAACTCATTCATCTGCTGCGCTTCGTTATCCGCGGCGGACGACC 447
 Qy 421 CGGCGGCAAGCTGCTCCCTTATATCTCATTCAGATGATCCGAGGCGCGGAGATT 480
 Db 448 CGTACCAAGCTGCTCCCTTATATCTCATTCAGATGATCCGAGGCGCGGAGATT 507
 Qy 481 AATCCATCTTTTGAAGGCTCGCCAAATACATTAACGCGGAGAGTCAATTTCTCCGAC 540
 Db 508 AATCCATCTTTTGAAGGCTCGCCAAATACATTAACGCGGAGAGTCAATTTCTCCGAC 567
 Qy 541 ATGTACATGCTGAGGTGAGACTAGTTGGGCGCAAAATCCAGCAAGTCCAGAGCT 600
 Db 568 GTGTACATCTGAGGTGAGACTAGTTGGGCGCAAAATCCAGCAAGTCCAGAGCT 627
 Qy 601 ACGGATGGCGTTTAAATAAACCCATTGGGTATATCCACCGTAACTTCGTGACG 660
 Db 628 ACGGATGGCGTTTAAATAAACCCATTGGGTATATCCACCGTAACTTCGTGACG 687
 Qy 661 TTGAGCAATGTTCCGAGCATGCGCAGCTTACGATCATGTTGTATGTATGAGGAC 720
 Db 668 TTGAGCAATGTTCCGAGCATGCGCAGCTTACGATCATGTTGTATGTATGAGGAC 747
 Qy 721 CGACCATCTTCCTC 734
 Db 748 CGGCATCTTCCTC 761

RESULT 14

AAZ09105
ID AAZ09105 standard; DNA; 768 BP.

XX AAZ09105;

DE 18-OCT-1999 (first entry)

DE Mistletoe lectin A2 DNA fragment.

KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A2; ds.

OS Viscum album.

XX DE19804210-A1.

XX 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.
 XX 03-FEB-1998; 98DE-1004210.
 XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX Morris P, Stiefel T, Voelter W, Welters P;
 PI WPI: 1999-445335/38.
 DR P-PSDB; AAY25981.

XX Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 PS Claim 15; Fig 3A; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC encodes a fragment of the mistletoe lectin A2 protein.

XX Sequence 768 BP; 175 A; 224 C; 182 G; 187 T; 0 other;

Query Match 73.4%; Score 614; DB 20; Length 768;
 Best Local Similarity 90.4%; Pred. No. 2.1e-190;
 Matches 669; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 1 CATCAGACGAGGCGGAGCAATATTTCGGTTCATCAGCTTCCGAGATTATGTCGA 60
 DB 28 CATCAACACAGGCGGAGTGAATACCTCCGGTTCATCAGCTTCCGAGATTATGTCGA 87
 QY 61 AGCGGAAGCTTTTCCAAATGAGATACACTTTGGCTCAGTCTAGATCCCGCTCGGAT 120
 DB 88 AGCGGAAGCTTTTCCAAATGAGATACACTTTGGCTCAGTCTAGATCCCGCTCGGAT 147
 QY 121 GCGCAAGATTTGTGTGTGGAATCTACCAATCAGGCGGAGACTGCATCAGCGCCG 180
 DB 148 GCGCAAGATTTGTGTGTGGAATCTACCAATCAGGCGGAGACTGCATCAGCGCCG 207
 QY 181 ATCGAGCTTACCAATCTGTAGGTGTGCTTACCAAGGAGCGACCAATCTACTTTTGG 240
 DB 208 ATCGAGCTTACCAATCTGTAGGTGTGCTTACCAAGGAGCGACCAATCTACTTTTGG 267
 QY 241 GCGGAGGAGCAAGAGGCGGCGGAAAGGATCTCTACCGGCAACAC-----AGATCC 294
 DB 268 GCGGAGGAGCAAGAGGCGGCGGAAAGGATCTCTACCGGCAACAC-----AGATCC 327
 QY 295 TCTCTCCATTCACCGGAGGTACACAGATCTGAGCGATCGCGGTCTATAGGAGCCAG 354
 DB 328 TCTCTCCATTCACCGGAGGTACACAGATCTGAGCGATCGCGGTCTATAGGAGCCAG 387
 QY 355 ATCCCTGTGGGTATAGAGAACTCATTTCAATCGTCTGCGGCTTCTGTTATCCAGCGG 414
 DB 388 ATCCCTGTGGGTATAGAGAACTCATTTCAATCGTCTGCGGCTTCTGTTATCCAGCGG 447
 QY 415 AGCAGCGGCGGCGGAGCTCGTTCCTTAAATCCATTCAGATGATGATCCGAGGCGCG 474
 DB 448 AGCAGCGGCGGCGGAGCTCGTTCCTTAAATCCATTCAGATGATGATCCGAGGCGCG 507
 QY 475 AGATTCAATCCCATTTTGGAGGCGGCGCAATACATTAACAGGCGGAGTCTATTTCTT 534
 DB 508 AGATTCAATCCCATTTTGGAGGCGGCGCAATACATTAACAGGCGGAGTCTATTTCTG 567

QY 535 CCCGACATGTACATGCTGAGCTGAGACTAGTTGGGCGCAATCCACGCAAGTCAG 594
 DB 568 CCAGACATGTACATGCTGAGCTGAGAGAGTGGGCGCAATCCACGCAAGTCAG 627
 QY 595 CAGTCAATGAGATGGCTTTTAAATACCATTTGGTGGGTATATCCACCGGTAACTTC 654
 DB 628 CATTCACCGATGGCTTTTAAATACCATTTGGTGGGTATATCCACCGGTAACTTC 687
 QY 655 GTGACGTTGAGCATGTTGCGGAGCATGCGGAGCTTACGATCATGTTGTATAGT 714
 DB 688 GTGACGTTGCTATATGTTGCTCTGTGATCGCAGCTTGGCATCATGTTGTATAGT 747
 QY 715 AGGAGCGGACCATCTTCTTC 734
 DB 748 GGAGAGCGGCGCATCTCTC 767

RESULT 15

AAC85474
 ID AAC85474 standard; cDNA; 768 BP.

XX AAC85474;
 AC AAC85474;

XX 16-MAY-2001 (first entry)

XX A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #3.

XX Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;

KW KML; tumour; KM-110; KML-C; KMBP; KML-IIU; KML-III;

KW heparin binding protein; ds.

XX Viscum album coloratum.

XX EP1074560-A2.

XX 07-FEB-2001.

XX 27-JUL-2000; 2000BP-0402168.

XX 27-JUL-1999; 99KR-0030638.

XX (MIST-) MISTLE BIOTECH CO LTD.

XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;

PI Kang T, Park C;

PI WPI: 2001-171044/18.

DR P-PSDB; AAB47092.

XX Novel lectin proteins isolated from Korean mistletoe, useful for

PT enhancing immunity and effectuating anti-tumoral activity -

PS Claim 5; Page 27-28; 62pp; English.

XX The sequences given in AAC85472-74 encode isoforms of an A-chain gene

CC involved in biosynthesis of lectins isolated from Korean mistletoe.

CC Korean mistletoe lectins (KML) are useful for enhancing immunity and

CC for treating tumours. The KML's are isolated from a protein fraction

CC derived from the leaves, stems and fruits of Korean mistletoe, which

CC is designated KM-110. One of the isolates, KML-C was shown to be

CC effective against colon 26-M3.1 carcinoma and L5178Y-ML25 lymphoma.

XX Sequence 768 BP; 171 A; 219 C; 200 G; 178 T; 0 other;

Query Match 70.3%; Score 588.6; DB 22; Length 768;
 Best Local Similarity 89.1%; Pred. No. 4.4e-182;
 Matches 662; Conservative 0; Mismatches 69; Indels 12; Gaps 2;

QY 1 CATCAGACGAGGCGGAGCAATATTTCGGTTCATCAGCTTCCGAGATTATGTCGA 60
 DB 28 CATCAACACAGGCGGAGTGAATACCTCCGGTTCATCAGCTTCCGAGATTATGTCGA 87
 QY 61 AGCGGAAGCTTTTCCAAATGAGATACACTTTCGCTCAGTCTACGATCCCGCTCGGAT 120

```
Db 88 |||||ACCGAAGCTTTCCAAATGACATACCGCTCCTGCTCGT---CAATCCGGTCTCCTCT 144
OY 121 GCGCAAGATTTGTGTGTGGAACTCACCAATC-----AGGGGGAGACTCGATC 171
Db 145 GCGGAGATTTGTGTGTGAACACAAATCAGTTGGGAAAGTGGAGAGACTCGATC 204
OY 172 ACGGGCCGATCGAGCTTACCAATCTGATGCTGTGGCTTACCAAGCAGGCAACATCC 231
Db 205 ACGGGCCGATCGAGCTTACCAATCTGATGCTGTGGCTTACCAAGCAGGCAACATCC 264
OY 232 TACTTTTGGCGGACGACACGAGCGGCGGAAAGGATCTTCAACGGGACACACAGA 291
Db 265 TACTTTTGGCGGACGACACGAGCGGCGGAAAGGATCTTCAACGGGACACACAGA 324
OY 292 TCCTCTCTCCCATTCACCGAGCTACACAGATCTGAGCGATAGCGCGGTATAGGAC 351
Db 325 TCCTCTCTCTTCATCAGGACGATCTGATCTGAGCGGTATAGGAC 384
OY 352 CAGATCCCTCTGGGTATAGAGACTCATTCATCCGCTCGGCGCTTATATCCAGGC 411
Db 385 CGGATCCCTCTGGGTATAGAGACTCATTCATCCGCTCGGCGCTTATATCCAGGC 444
OY 412 GCGAGCACCGGCGCCCAAGCTGCTTCCCTATATCTCTCATTCAGATCTCCGAGGC 471
Db 445 GCGAGCACCGGCGCCCAAGCTGCTTCCCTATATCTCTCATTCAGATCTCCGAGGC 504
OY 472 GCGAGATTCATCCATCTTTTGGAGGCTCGCAATACATTAACAGGCGGAGTCATTT 531
Db 505 GCCAGATTCATCCATCCCTATGAGAGGCTCGCAATACATTAACAGTGGGTGTCATAT 564
OY 532 CTTCCCGACATGTACATGCTGAGAGCTGGAGACTAGTTGGGCGCAACAATCCAGCAAGTC 591
Db 565 CTTCCAGAGCTGTACATGCTGAGAGCTGGAGAGTGGGCGCAACAATCCAGCAAGTC 624
OY 592 CAGCACTCTACGAGATGGCGTTTAAATAACCAATTCGTTGGGTATATCCAGCGTAAC 651
Db 625 CAGCACTCTACGAGATGGCGTTTAAATAACCAATTCGTTGGGTATATCCAGCGTAAC 684
OY 652 TTCGTGACGTTGAGCAATGTCGAGAGTGAATCGCAGCTTAGCAGATCATGTTGTGA 711
Db 685 TTCGTGTTGTTGAGCAATGTCGAGAGTGAATCGCAGCTTAGGAGATCATGTTGTGA 744
OY 712 TGTAGGACCGACCATCTTCTC 734
Db 745 TGCAGGACCGGTATCTTCCC 767
```

Search completed: March 26, 2003, 18:18:10
Job time : 236.832 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 18:18:22 ; Search time 50.4724 Seconds

(without alignments)
5085.723 Million cell updates/sec

Title: US-09-627-165E-15

Sequence: 1 catcagacgacgagcgagcagca.....aacccacgctgcacgtcga 837

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCrUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	695.2	83.1	1923	4 US-08-776-059-34	Sequence 34, Appl
2	645.2	77.1	774	4 US-08-776-059-30	Sequence 30, Appl
3	628.2	75.1	705	4 US-08-776-059-38	Sequence 38, Appl
4	78.8	9.4	1140	1 US-08-218-303-15	Sequence 15, Appl
5	78.8	9.4	1140	2 US-08-338-793D-60	Sequence 60, Appl
6	77.2	9.2	807	2 US-08-356-786-7	Sequence 7, Appl
7	77.2	9.2	1605	2 US-08-356-786-9	Sequence 9, Appl
8	77.2	9.2	1855	4 US-09-147-208-33	Sequence 33, Appl
9	77.2	9.2	1855	4 US-09-147-208-40	Sequence 40, Appl
10	77.2	9.2	1855	4 US-09-147-208-47	Sequence 47, Appl
11	77.2	9.2	1855	4 US-09-147-208-54	Sequence 54, Appl
12	77.2	9.2	1879	4 US-09-147-208-23	Sequence 23, Appl
13	77.2	9.2	1879	4 US-09-147-208-25	Sequence 25, Appl
14	77.2	9.2	1879	4 US-09-147-208-25	Sequence 25, Appl
15	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, Appl
16	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, Appl
17	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, Appl
18	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, Appl
19	66.4	7.9	813	2 US-08-646-360-11	Sequence 11, Appl
20	66.4	7.9	813	2 US-08-646-360-11	Sequence 11, Appl
21	66.4	7.9	813	3 US-08-839-765-11	Sequence 24, Appl
22	66.4	7.9	813	3 US-08-839-765-11	Sequence 24, Appl
23	66.4	7.9	813	3 US-09-136-389-11	Sequence 11, Appl
24	66.4	7.9	813	4 US-09-217-352-26	Sequence 24, Appl
25	65	7.8	774	5 US-09-610-838-11	Sequence 11, Appl
26	65	7.8	955	4 US-08-621-803-258	Sequence 258, Appl
27	65	7.8	1003	2 US-08-621-803-258	Sequence 252, Appl

28	65	7.8	1003	4 US-09-217-352-252	Sequence 252, App
29	65	7.8	1072	2 US-08-621-803-250	Sequence 250, App
30	65	7.8	1072	2 US-09-217-352-250	Sequence 250, App
31	63.2	7.6	813	1 US-07-901-707-11	Sequence 11, Appl
32	63.2	7.6	813	1 US-07-901-707-11	Sequence 11, Appl
33	63.2	7.6	813	1 US-07-988-430-57	Sequence 57, Appl
34	63.2	7.6	813	1 US-07-988-430-57	Sequence 57, Appl
35	63.2	7.6	813	5 PCT-US92-09487-11	Sequence 11, Appl
36	63.2	7.6	813	5 PCT-US92-09487-11	Sequence 11, Appl
37	59	7.0	913	1 US-07-901-707-13	Sequence 57, Appl
38	59	7.0	913	1 US-07-988-430-13	Sequence 13, Appl
39	59	7.0	913	1 US-08-425-336-13	Sequence 13, Appl
40	59	7.0	913	1 US-08-425-336-13	Sequence 13, Appl
41	59	7.0	913	1 US-08-425-336-13	Sequence 13, Appl
42	59	7.0	913	1 US-08-425-336-13	Sequence 13, Appl
43	59	7.0	913	2 US-08-646-360-13	Sequence 13, Appl
44	59	7.0	913	3 US-08-839-765-13	Sequence 13, Appl
45	59	7.0	913	3 US-09-136-389-13	Sequence 13, Appl
				4 US-09-610-838-13	Sequence 13, Appl

ALIGNMENTS

```

RESULT 1
US-08-776-059-34
Sequence 34, Application US/08776059B
Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LENTZEN, Hans
APPLICANT: ECK, Jurgen
APPLICANT: BAUR, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-2003
CURRENT APPLICATION NUMBER: US/08/776,059B
EARLIER FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 9510949.8
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 1923
TYPE: DNA
ORGANISM: Viscum album
US-08-776-059-34

Query Match
Best Local Similarity 83.1%; Score 695.2; DB 4; Length 1923;
Matches 758; Conservative 0; Mismatches 63; Indels 15; Gaps 1;

QY 1 CATCAGACGAGCGGAGAGATATTTCCGGTTCACGCTTCGAGATATGTCCTA 60
    |||||
DB 181 CATCAACACGAGCGGAGAGATATTTCCGGTTCACGCTTCGAGATATGTCCTA 240
    |||||
QY 61 AGCGAAGCTTTTCATGATGATATTCCTTGGTCTAGTCTACATCCCGTCTGGAT 120
    |||||
DB 241 AGCGAAGCTTTTCATGATGATATTCCTTGGTCTAGTCTACATCCCGTCTGGAT 300
    |||||
QY 121 GCGCAAGATTTGTTGTGTGAGAACCTACCAATCAGGGGAGACATCAGCGCGCC 180
    |||||
DB 301 GCGCAAGATTTGTTGTGTGAGAACCTACCAATCAGGGGAGACATCAGCGCGCC 360
    |||||
QY 181 ATCAGCTTACCAATCTGTACGTGCTTACCAAGGAGGACCAATCCTACTTTTG 240
    |||||
DB 361 ATCAGCTTACCAATCTGTACGTGCTTACCAAGGAGGACCAATCCTACTTTTG 420
    |||||
QY 241 CCGGAGCAGCAACGAGCGGAGGAGGATCTTTCACCGGAGCAGCAATCCTCTCTC 300
    |||||
DB 421 CCGGAGCAGCAACGAGCGGAGGAGGATCTTTCACCGGAGCAGCAATCCTCTCTC 480
    |||||
QY 301 CCATTTCACCGGAGGAGGATCTTTCACCGGAGGAGGATCTTTCACCGGAGGAGGATCTTCTC 360
    |||||

```

Db	481	CCATTACGAGAACGTAACCTTATCTGTGGAGGATACGCCGGACATATAGGGACCAAGATCCTT	540
OY	361	CTGGGTTATAGGAATCTCATTAATCCGTCTCCGGCGCTTGTTATTCAGAGCGGACAGCAC	420
OY	541	CTCGGTATTAACCAACTCATTAATACCGGACAGGGCGCTTGCTTTCCGGGCGGACAGCAG	600
Db	421	CGGGCCCCAGCTCGTTCCCTTATTAATCTCATTCAGATGATCTCCAGGCGCGAGATTC	480
OY	601	CGTACCCAGACGTCTTCGATTTTAATCCTCATTCAGATGATCTCCAGGCGCGAGATTC	660
Db	481	AATCCCAATCTTTGGAGGGCTGGCCAAATTCATTAACAGCGGGGAGTCATTTCTCCGAC	540
OY	661	AATCCCAATCTTTATGAGAGGCTGGCCAAATTCATTAACAGTGGGGCGTCATTTCTCCGAC	720
Db	541	ATGTACATGCTCGAGCTGGAGACTAGTTGGGGCCAAACATTCAGCGAAAGTCAGAGCTGT	600
OY	721	GTTGTACATGCTGGAGCTGGAGAGCAATTTGGGGCCAAACATTCAGCGAAAGTCAGAGCTGT	780
Db	601	ACGAGTGGCGTTTATTAACCAATTCGTTGGGTATATCCACCGGTAACTTTCGTAGC	660
OY	781	ACGAGTGGCGTTTATTAACCAATTCGTTGGGTATATCCACCGGTAACTTTCGTAGC	840
Db	661	TTGAGCAATGTTCCGACGATATGCCACCTTAAGGATCATAGTTGTTGTAATAGGAGAC	720
OY	841	TTGAGCAATGTTCCGACGATATGCCACCTTAAGGATCATAGTTGTTGTAATAGGAGAG	900
Db	721	CGACCAATCTTCCTCGAGAGTGGCGGTATTTGGCGCTGGTCATACGACCCGCTT	780
OY	901	CGACCAATCTTCCTCGAGAGTGGCGGTATTTGGCGCTGGTCATACGACCCGCTAT	960
Db	781	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	840
OY	961	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1020
Db	841	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	900
OY	1021	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1080
Db	901	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	960
OY	1081	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1140
Db	961	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1020
OY	1141	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1200
Db	1021	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1080
OY	1201	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1260
Db	1081	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1140
OY	1261	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1320
Db	1141	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1200
OY	1321	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1380
Db	1201	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1260
OY	1381	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1440
Db	1261	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1320
OY	1441	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1500
Db	1321	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1380
OY	1501	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1560
Db	1381	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1440
OY	1561	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1620
Db	1441	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1500
OY	1621	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1680
Db	1501	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1560
OY	1681	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1740
Db	1561	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1620
OY	1741	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1800
Db	1621	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1680
OY	1801	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1860
Db	1681	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1740
OY	1861	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1920
Db	1741	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1800
OY	1921	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1980
Db	1801	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1860
OY	1981	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2040
Db	1861	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1920
OY	2041	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2100
Db	1921	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	1980
OY	2101	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2160
Db	1981	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2040
OY	2161	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2220
Db	2041	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2100
OY	2221	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2280
Db	2101	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2160
OY	2281	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	2340
Db	2161	AGGCGGCGCGTGGAGATTTACCTGCATCGATCGATCTTCGAAACCCACCGTCGATCGT	222

```

RESULT 2
US-08-776-059-30
: Sequence 30, Application US/08776059B
: Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LENTZEN, Hans
APPLICANT: ECK, Jurgen
APPLICANT: BAUR, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-2003
CURRENT APPLICATION NUMBER: US/08/776, 059B
CURRENT FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 95109949.8
EARLIER FILING DATE: 1995-06-26
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 30
: LENGTH: 774
: TYPE: DNA
: ORGANISM: Viscum album
US-08-776-059-30

Query Match          77.1%; Score 645.2; DB 4; Length 774;
Best Local Similarity 92.7%; Pred. No. 1.2e-207;
Matches 677; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

DB

61 AGCGAAGCTTTCCAAATGACATACTTTCGTGCATGCTTAGCATCCCCTCCGAT 120
CY

94 AGGGGAAGCTTTCCCAATGACATACACTCTTGCGTCATGCTAACGATCCCCGTCGAT 153
DB

QY	121	GGCCAAAGATTTGTGTGGTGAAC	CTCACCAATCGAGGGGGAGAGT	GATGATACGGCGCC	180
Db	154	GGCCAAAGATTTGTGTGGTGAAC	CTCACCAAGGGGGAGAGT	GATGATACGGCGCC	213
QY	181	ATGAGCGTTACCATCTGTACGTG	TGAGTGTACCAAGCAGCGCAAT	TCTACTTTTG	240
Db	214	ATGAGCGTTACCATCTGTACGTG	TGAGTGTACCAAGCAGCGCAAT	TCTACTTTTG	273
QY	241	CGGAGCGACCAAGGGCGGGAAG	GCATCTCTTACCGGCAACCAAGAT	CTCTCTC	300
Db	274	CGGAGCGACCAAGGGCGGGAAG	GCATCTCTTACCGGCAACCAAGAT	CTCTCTC	333
QY	301	CCATTCCACGGGAAGCTACACAG	ATCTTGAGCGATACCGCGTCA	TAGGAGACAGTCCCT	360
Db	334	CCATTCCACGGGAAGCTACCTGT	ATCTTGAGCGATACCGCGATAG	GGAGACAGTCCCT	393
QY	361	CTGGGTATAGGGAACCATTCAT	TCAATCGTCTCGGCGCTTCT	GTATCCAGGCGGACAC	420
Db	394	CTCGGTATAGGGAACCATTCAT	TCAATCGTCTCGGCGCTTCT	GTATCCAGGCGGACAC	453
QY	421	CGGGCCCAAGCTGCTTCCCTTA	TAAATCCATTCAGATGATCTCC	AGGACCGCGAGATTC	480
Db	454	CGTACCCAAAGCTGCTGATTTA	TATCTCTTATTCAGATGATCT	CGAGGCGCGCGAGATTC	513
QY	481	AATCCCATCTTTTGGAGGGCTC	GCCTAATCATTTAACAAGCGGG	AGATCATTTCTCCGAC	540
Db	514	AATCCCATCTTTATGAGGGGCT	CGCCATTAATTTAACAGTGGG	CGATATTTCTGCCAGAC	573
QY	541	ATGTCATAGTCTGAGCTGGAAG	TAGTTGGGGCCAAATTCACG	AGATCCAGAGTCT	600
Db	574	GTGTACATCTGTGAGCTGTGA	ACAGTTGGGGCCAAATTCACG	AGATCCAGAGTCT	633
QY	601	ACGAGTGGCGTTTAAATAACCA	ATTCCGTTGGGTATTCACACGG	TACTCTCGAGAC	660
Db	634	ACGAGTGGCGTTTAAATAACCA	ATTCCGTTGGGTATTCACACGG	TACTCTCGAGAC	693
QY	661	TTGAGCAATGTTGGGAGCGTGA	TGATGCCACGCTTAGCATCAT	GTGTTGTATATAGGAG	720
Db	694	TTGAGCAATGTTGGGAGCGTGA	TGATGCCACGCTTAGCATCAT	GTGTTGTATATAGGAG	753
QY	721	CGACCATCTT	730		
Db	754	CGGCAATCTT	763		

```

RESULT 3
US-08-776-059-38
; Sequence 38, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776.059B
; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Viscum album
US-08-776-059-38

```

Query Match	75.1%;	Score 628.2;	DB 4;	Length 705;
Best local Similarity	93.2%;	Pred. No. 6.4e-202;		
Matches 657; Conservative	0;	Mismatches 48;	Indels 0;	Gaps 0;

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:25:55 ; Search time 6.57789 Seconds
(without alignments)
833.848 Million cell updates/sec

Title: US-09-627-165E-14
Perfect score: 596
Sequence: 1 ARPNPIKWRRLRQINSSESS.....VVDANDVTCTXSEPTVATV 119

Scoring table: PROSUM62
Gapop 10.0, Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications MA:

1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEM_PUB pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB pep: *
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7: /cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep: *
10: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB pep: *
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep: *
12: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB pep: *
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	311	52.2	252	10 US-09-347-064-2	Sequence 2, Appl1
2	311	52.2	252	10 US-09-347-064-8	Sequence 8, Appl1
3	772	12.1	247	9 US-09-792-793A-39	Sequence 39, Appl1
4	70.5	11.8	293	10 US-09-765-527-259	Sequence 259, App
5	70.5	11.8	332	10 US-09-765-527-253	Sequence 253, App
6	69.5	11.7	309	10 US-09-765-527-253	Sequence 253, App
7	69.5	11.7	477	9 US-09-738-626-645	Sequence 645, App
8	69	11.6	247	9 US-09-792-793A-34	Sequence 34, Appl1
9	68.5	11.5	251	10 US-09-765-527-247	Sequence 247, App
10	68.5	11.5	1018	10 US-09-801-574-32	Sequence 32, Appl1
11	65.5	11.0	937	10 US-09-949-197-19	Sequence 19, Appl1
12	65	10.9	1110	9 US-10-149-819-12	Sequence 12, Appl1
13	64.5	10.8	951	9 US-09-924-097-15	Sequence 15, Appl1
14	64	10.7	263	10 US-09-347-064-10	Sequence 10, Appl1
15	64	10.7	267	10 US-09-347-064-4	Sequence 4, Appl1
16	62.5	10.5	136	9 US-10-020-419-4	Sequence 4, Appl1
17	62	10.4	366	10 US-09-764-864-808	Sequence 808, App
18	61	10.2	1229	10 US-09-815-242-5818	Sequence 5818, App
19	61	10.2	1229	10 US-09-815-242-12946	Sequence 12946, A

20	60.5	10.2	260	10 US-09-742-693-31	Sequence 31, Appl1
21	60	10.1	115	10 US-09-746-801A-47	Sequence 47, Appl1
22	60	10.1	144	10 US-09-864-761-37049	Sequence 37049, A
23	60	10.1	211	10 US-09-728-911-34	Sequence 34, Appl1
24	60	10.1	212	9 US-09-912-672A-6	Sequence 6, Appl1
25	60	10.1	240	10 US-09-864-761-47947	Sequence 47947, A
26	60	10.1	312	10 US-09-912-672A-5	Sequence 5, Appl1
27	60	10.1	560	9 US-09-912-672A-2	Sequence 2, Appl1
28	60	10.1	574	9 US-09-864-761-164	Sequence 164, App
29	60	10.1	574	9 US-10-063-547-164	Sequence 164, App
30	60	10.1	574	9 US-10-063-547-164	Sequence 164, App
31	60	10.1	574	9 US-10-063-502-164	Sequence 25, Appl1
32	60	10.1	574	10 US-09-728-911-25	Sequence 4, Appl1
33	60	10.1	574	10 US-09-870-571-4	Sequence 164, App
34	60	10.1	574	10 US-10-006-867-164	Sequence 33522, A
35	60	10.1	777	10 US-09-864-761-33522	Sequence 2, Appl1
36	59.5	10.0	552	10 US-09-835-654-2	Sequence 522, App
37	59	9.9	110	10 US-09-925-302-522	Sequence 4, Appl1
38	59	9.9	262	10 US-09-978-274A-4	Sequence 11800, A
39	59	9.9	263	10 US-09-978-274A-4	Sequence 2, Appl1
40	59	9.9	268	10 US-09-815-242-11800	Sequence 35, Appl1
41	59	9.9	314	10 US-09-978-274A-2	Sequence 44, Appl1
42	58.5	9.8	275	9 US-09-801-368-44	Sequence 14, Appl1
43	58.5	9.8	876	10 US-09-738-626-4478	
44	58	9.7	413	9 US-09-479-614-14	
45	58	9.7	431	9 US-09-479-614-14	

ALIGNMENTS

RESULT 1
US-09-347-064-2
Sequence 2, Application US/09347064A
Patent No. US20020045208A1
GENERAL INFORMATION:
APPLICANT: Eck, Jürgen
APPLICANT: Schmidt, Arno
TITLE OF INVENTION: Holger
TITLE OF INVENTION: Ribosome-inactivating Proteins of the mistletoe Viscum
TITLE OF INVENTION: album
FILE REFERENCE: 09282-5
CURRENT APPLICATION NUMBER: US/09/347, 064A
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012-0
EARLIER FILING DATE: 1997-01-02
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 252
TYPE: PRT
ORGANISM: Viscum album
US-09-347-064-2

Query Match 52.2% Score 311: DB 10: Length 252:
Best local similarity 71.8% Pred. No. 2.6e-30;
Matches 61: Conservative 7: Mismatches 17: Indels 0: Gaps 0:

OY 1 ARPNPIKWRRLRQINSSESSPPNNYMLELTSNGRSTGVQOOSKDCIFNTQRLQISAGN 60
DB 168 ARPNPIKWRRLRQINSSESSPPNNYMLELTSNGRSTGVQOOSKDCIFNTQRLQISAGN 227
OY 61 FVTXSNVRDVVISLAIMLFECCGRP 85
DB 228 FVTXSNVRDVVISLAIMLFECCGRP 252

RESULT 2
US-09-347-064-8
Sequence 8, Application US/09347064A

Patent No. US20020045208A1
GENERAL INFORMATION:
APPLICANT: Eck, Jurgen
APPLICANT: Schmidt, Arno
APPLICANT: Zinke, Holger
TITLE OF INVENTION: Recombinant Fusion Proteins Based on
TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
FILE REFERENCE: 09282-5
CURRENT APPLICATION NUMBER: US/09/347,064A
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012.0
EARLIER FILING DATE: 1997-01-02
SOFTWARE: Patentin Ver. 2.1
SEQUENCE ID NO: 8
LENGTH: 252
TYPE: PRF
ORGANISM: Viscum album
US-09-347-064-8

Query Match 52.2%; Score 311; DB 10; Length 252;
Best Local Similarity 71.8%; Pred. No. 2,6e-30;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;
QY 1 ARENPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 60
DB 167 ARFNPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 60
QY 61 EYTXSNVNDVYISLAIMLFECSSGRF 85
DB 227 EYTLFNVNDVYISLAIMLFECSSGRF 251

RESULT 3
US-09-792-793A-39
Sequence 39, Application US/09792793A
Patent No. US20020168370A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Coggin, Philip
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
FILE REFERENCE: 25020-601D
CURRENT APPLICATION NUMBER: US/09/792,793A
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 93
SOFTWARE: Patentin Ver. 2.0
SEQUENCE ID NO: 3
LENGTH: 24
TYPE: PRF
ORGANISM: Trichosantheus kirilowii
US-09-792-793A-39

Query Match 12.1%; Score 72; DB 9; Length 247;
Best Local Similarity 26.8%; Pred. No. 0.46;
Matches 22; Conservative 24; Mismatches 30; Indels 6; Gaps 3;
QY 1 ARENPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 58
DB 162 ARFNPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 58
QY 59 GNVYYSNVNDVYISLAIMLFECSSGRF 78
DB 220 GNVYYSNVNDVYISLAIMLFECSSGRF 219

RESULT 4
US-09-765-527-259
Sequence 259, Application US/09765527
Patent No. US20020006638A1

GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Roole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 259:
SEQUENCE CHARACTERISTICS:
LENGTH: 293 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 259:
US-09-765-527-259

Query Match 11.8%; Score 70.5; DB 10; Length 293;
Best Local Similarity 26.6%; Pred. No. 0.89;
Matches 25; Conservative 19; Mismatches 47; Indels 3; Gaps 2;
QY 1 ARENPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 59
DB 190 ARFNPXWRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDFNFIQIRLQISAGN 59
QY 60 NEVYYSNVNDVYISLAIMLFECSSGRFSSLDHPS 93
DB 248 KRYVTAVDQVKPKIALKLFVKKPKSALDPKS 281

RESULT 5
US-09-765-527-251
Sequence 251, Application US/09765527
Patent No. US20020006638A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSER: Marshall, O'Roole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 5

US-08-338-793D-60

; Sequence 60, Application US/08338793D

; Patent No. 5840521

; GENERAL INFORMATION:

; APPLICANT: Barth, Peter Thomas

; TITLE OF INVENTION: VECTOR

; NUMBER OF SEQUENCES: 61

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: CUSHMAN DABRY CUSHMAN

; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF

; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.

; STREET: 1100 New York Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3918

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage

; COMPUTER: IBM PC/XT/AT Compatibles

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Microsoft word or ASCII editors

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/338,793D

; FILING DATE: 08-No. 5840521-94

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 07/842,081

; FILING DATE: 26-Feb-92

; CLASSIFICATION: 435

; APPLICATION NUMBER: 9104017.0

; FILING DATE: 26-Feb-91

; APPLICATION NUMBER: 9109188.4

; FILING DATE: 29-Apr-91

; ATTORNEY/AGENT INFORMATION:

; NAME: Kokulis, Paul N.

; REGISTRATION NUMBER: 16,773

; REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-861-3000

; TELEFAX: 202-822-0944

; TELEFAX: 6714627 CUSH

; INFORMATION FOR SEQ ID NO: 60:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1140

; TYPE: Nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

; US-08-338-793D-60

; Query Match

; Best Local Similarity 9.4%; Score 78.8; DB 2; Length 1140;

; Matches 311; Conservative 0; Mismatches 297; Indels 24; Gaps 3;

; DB 124 CAAGATTTGTGTGTAAGTCAACATCAGGAGGAGAGTGTATCAGCGCCATC 183

; DB 305 CAACGTTTATTTAGTGAAGTCAATCATGACGAGCTTCTGTTACATTAAGCTG 364

; DB 184 GACGTTACCAATCTGTAGCTGTGCTTACCAAGCAGGCGAACCAATCTTCTTTTGGG- 242

; DB 365 GATGACCAATGATATGTGTGCTGCTACCGTGTGAATAAGGCATATTTCTTTTAT 424

; DB 243 -----CGAGCAGCAGAGCGCGCGGAAAGCATCTCTTACAGTGTCAAAATCGA 291

; DB 425 CCTGCAATCAAGAGATGAGAGCAATCACTCTTTTACAGTGTCAAAATCGA 484

; DB 292 TCCCTCTCCCAATCAGGAGGAGTACACAGATCTGAGGCGATAGCGGGTCAAT---AGG 348

; DB 485 TATACATTCGCTTGTGGTGATATATATGACTGTAACACTTGTCTGTATCTGGA 544

; DB 349 GACCAATCCCTCTGTGTATAGAGCACTCAATCCGCTCTGCGCCTTCTGTA---- 404

Db 545 GAAATATGAGTTGGGAAATGCTCCAGTACAGAGAGCTATCTCAGCGCTTATATTAC 604

QY 405 -----TCCAGGCGGACAGACCCGCGCCCAAGCTCGTTCCTTATATCTCATTTAGATG 459

Db 605 AGTACTGTGTGACACTGAGCTTCCAACTCTGGCTGCTTCTTATATATTTGATCCAAAAG 664

QY 460 ATCTCCAGGCGCGGAGATTCATTCCTTCTTGGAGGCGCTCGCCAAATTAACAGC 519

Db 665 ATTTCAAGAGCAGCAAGATTCATATATTGAGGAGGAAATGCGCAGCAATTAAGTATC 724

QY 520 GGGAGTCAATTTCTCCGACATGTACATGCTGAGCTGAGACTAGTTGGGCGCAACA 579

Db 725 AACCGAGATCTGACACAGATCTTACGTAATTAATTAATTAATTAATTAATTAATTAAT 784

QY 580 TCCAGCAATCCAGCAGCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 639

Db 785 TCCACTGCAATTCAGAGTCTAACCAAGAGCCCTTGTAGTCAATTAATTAATTAATTAATTAAT 844

QY 640 TCCAGCGTAACTTCGAGCTTGAAGCAATGTTCCGAGCTGATCCAGCTTATGAGATC 699

Db 845 CGTAAATGTTCCAAATTCATGCTGTAAGTATGATATTAATTAATTAATTAATTAATTAATTAAT 904

QY 700 ATGTTGTTGTATGTAGGAGCAGCAATCTTC 731

Db 905 ATGTTGTTGTATGTAGGAGCAGCAATCTTC 936

RESULT 6

US-08-356-786-7

; Sequence 7, Application US/08356786

; Patent No. 5877305

; GENERAL INFORMATION:

; APPLICANT: Huston, James S.

; APPLICANT: Oppermann, Hermann

; APPLICANT: Houston, L. L.

; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault

; STREET: Exchange Place, 53 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/356,786

; FILING DATE:

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 07/831,967

; FILING DATE: 06-Feb-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Pitcher, Edmund R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CRP-053

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 807 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

```

;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..807
;      OTHER INFORMATION:
US-08-356-786-7      /note= "product = "Riclin-A chain"

```

Query Match	9.28;	Score 77.2;	DB 2;	Length 807;
Best Local Similarity	49.18;	Pred. No. 3.8e-16;		
Matches 310; Conservative	0;	Mismatches 298;	Indels 24;	Gaps 3

QY	124	CAAGATTTTGTGGTGGGAACACCAATCAGAGGGGGAACATCGATCAGCGGCGATC	183
Db	166	CAACGGTTATTTAGTTGAACCTCQAATCATGCAAGACTTCTGTACATTAGCGCTG	225
QY	184	GAGCTTACCAATCTGTACGTGTGGCTTACCAAGCAGGCGCAACATCTACTTTTGGC-	242
Db	226	GATGTACCAATGCATATGTGTGATAGGCTACCGTGTGGAATAGCGCATTTTCTTTCAT	285
QY	243	-----CGAGCCAGCAGAGCGGGCGGAAAGCATCTCTTACCGGACACACAGAGA	291
Db	286	CCTGACAAATCAGGAAGATGCGAAGCAATCATCATCTTTTCACTGATATGTTCAAAATCGA	345
QY	292	TTCCTCTTCCCATTTACCCGGAAGCTACACAGATCTGAGCGATACGCCGCTAT---	Agg 348
Db	346	TATACATTCGCTTGGTGTATTTATGATATGAACTTGAAACACTTGTGTGTAATCTGAGA	405
QY	349	GACCAATCCCTCTGGGTATAGAGAACTATTCATCCGCTCGGCGCTTCGTTA----	404
Db	406	GAATATATCAGATTGGGAATGTGCCACTAGAGAGGCTATCGAGGCTTTATTATTC	465
QY	405	-----TCCAGGCGGCGACACCACCGGCCAAGCTCGTCTTCTTAAATCCATTTACATG	459
Db	466	AGTACTGTGTGCACTCAGCTTCCAACTCTGGCTGCTCTTAAATTTGCATTCGCAAAAT	525
QY	460	ATCTCGAGGCCCGAGATTCATTCCTATCTTTTGGAGGCTCGCCAAATCATTTAACAGC	519
Db	526	ATTTCAGAACAGCAGAAATATTCCAATATTTAGAGGAGAAATGTGGCGAAGAAATTAGTAC	585
QY	520	GGGAGTCATTTCTCCGCAATGTACATCTCGAGGTGAGACTGATGGGGCCAAACA	579
Db	586	AACCGAGATCTGCACCAAGTCTCTAGCGTAAATTACCTTAGAATATGTGGGAGAGCTT	645
QY	580	TCCACGCAAGTCCAGCAGTCTACGGATGCGGTTTTTAATAACCATTTTGGTGGGTATA	639
Db	646	TCCACTGCAATTCMAAGATCTAACCAAGAGCGCTTTGCTGATCCAAATTCMACTGCAAGA	705
QY	640	TCCACGCGTAATCTCGAGAGTGTAGACAAATGTGCGACATGATGCGACACTTAGAGAT	699
Db	706	CGTAATGTTTCCAATTCAGTGTGTAGAGATGTAGTATATTATTCCTATCATAGCTCTC	765
QY	700	ATGTTGTTTGTATGTAGGAGACGACCATCTTC	731
Db	766	ATGGTGTATGATGCGACCTTCACCATCTGTC	797

RESULT 7
 US-08-356-786-9
 : Sequence 9, Application US/08356786
 : Patient No. 5877305
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Huston, James S.
 : APPLICANT: Oppermann, Hermann
 : APPLICANT: Houston, L. L.
 : APPLICANT: Ring, David B.
 : TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
 : TITLE OF INVENTION: Marker
 : NUMBER OF SEQUENCES: 16
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Edmund R. Pletcher, Teesta, Hurwitz, & Thibeault
 : STREET: Exchange Place, 53 State Street
 : CITY: Boston
 : STATE: Massachusetts
 : COUNTRY: USA

```

1 ZIP: 02109
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: Floppy disk
5 COMPUTER: IBM PC compatible
6 OPERATING SYSTEM: PC-DOS/MS-DOS
7 SOFTWARE: PatentIn Release #1.0, Version #1.2
8 CURRENT APPLICATION DATA:
9 APPLICATION NUMBER: US/08/356,786
10 FILING DATE:
11 CLASSIFICATION: 424
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: 07/831,967
14 FILING DATE: 06-FEB-1992
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Pletcher, Edmund R.
17 REGISTRATION NUMBER: 27,829
18 REFERENCE/DOCKET NUMBER: CRP-053
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (617) 248-7000
21 TELEFAX: (617) 248-7100
22 INFORMATION FOR SEQ ID NO: 9:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 1605 base pairs
25 TYPE: nucleic acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28 MOLECULE TYPE: DNA (genomic)
29 FEATURE:
30 NAME/KEY: CDS
31 LOCATION: 1..1605
32 OTHER INFORMATION: /note= "product = "G-FTT""
33
34 US-08-356-786-9

```

Query Match	9.2%	Score 77.2;	DB 2;	Length 1605;
Best Local Similarity	49.1%;	Pred. No. 5.4e-16;		
Matches 310;	Conservative	0;	Mismatches 298;	Indels 24; Gaps 3

QY	124	CAAAAGTTTGTGTGTGTGTGAGCTCACCACATCAGGGGGAGACTCGTCACGGCCCATC	183
Db	172	CAACGGTTATTATTAGTTGAACTCTCAAAATCATGACAGACTTTCGTTCATATAGCGCTG	233
QY	184	GACGTTACCAATCTGTACGTGTGGCTTACCAGACAGCGACCACAAATCCACTATTTTGGC-	242
QY	243	-----CGAGCGACCGAGCGGGGGGAAAGGCATCTGTACCGGCACCAACAGA	291
Db	232	GATGTCAACCAATGCAATATGTAGGTAGCGCTACGGCTGTGGAAATAGGCATATTTCTTTCAT	291
QY	292	CCTGCACATCAGGAAGATGCGAAAGCAATACATCTTTTACGTGATGTTCAAAATCGA	351
Db	352	TATACATTTGCGCTTGTGGTAAATTAATGATAGACTTGAACTGATGCTGATATCTGAGA	411
QY	292	TCTCTCTCCCAATTCACGGAGCTACAGATCTGTGAGCGATACCGCGTCAAT--AGC	348
Db	352	TATACATTTGCGCTTGTGGTAAATTAATGATAGACTTGAACTGATGCTGATATCTGAGA	411
QY	349	GACCAATCCCTCCTGGGATATAGAGAGACATTCATCCGCTCGCGGCTTCGTTA----	404
Db	412	GAAATATGACTTGGGAAATGTCTCACTAGAGAGAGCTATCTCAGCGCTTATTTATTATAC	471
QY	405	-----TCCAGCGCGAGACACCGGGGCCAAGCTGTTCCTTATATATCCTCATTCAGATG	459
Db	472	AGTACTGTGGCACTCAGCTTCCTCAACTGTGGCGTTCCTTATATTTGGATCCCAATG	531
QY	460	ATCTCCGAGCGCCGCAATTTCAATCCCATCTTTTGGAGGGCTCGCCAAATTTATACAGC	519
Db	532	ATTTCAGAAAGCAGCAAGATTTCAATATATTATAGGAGGAAATCGCAGCAGAAATTAAGTAC	591
QY	520	GGGAGATCTATTTCTTCCCGACATGTACATCTCTGAGCTGGAGACTATGTGGGCCAACAA	579
Db	592	AACCGAGATCTGCACACAGATCTTAGGTAAATTACACTTGGAGATATAGTTGGGGAGACTT	651
QY	580	TCCACGACAGTCCAGCACTCTACGAGATGGCGTTTAAATACCCCAATTCGTTGGTGATATA	639
Db	652	TCCATGCAATTCACAGACTCTTACCAAGAGGACCTTGTGATGTCATCAATTCACATCAAGA	711


```

QY 184 GACGTTACCAATCTGTACGTGTGGCTTACCAAGCAGGACCAATCTCACTTTTTCG- 242
    || || || || || || || || || || || || || || || || || || || || ||
Db 334 GATGTCAACATGATATGTGTGGCTACCGCTGTGGAATAGCGATATTTCTTTCAT 393
QY 243 -----CGACGACCAAGACGCGCGGAAAGCATCTTTCACCGGACCAACCA 291
    || || || || || || || || || || || || || || || || || || || || ||
Db 394 CCTGACATCAGGAAGATGAGAGCAATCATCTCTTTTCACTGATGTTCAAAATCGA 453
QY 292 TCCTCTCCCATTCACCGGAGCTACAGANCTGGAGGAGATACGCGGCTCAT--AGG 348
    || || || || || || || || || || || || || || || || || || || || ||
Db 454 TATCATATTCGCTTTGGTGTGTAATTAATGATAGACTTGAACAATCTGCTGTAATCTGAGA 513
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCATTCATCCGTCGTGCGGCTTCGTTA--- 404
    || || || || || || || || || || || || || || || || || || || || ||
Db 514 GAAATATCGAGTTGGGAAATGTGTCACATAGAGAGGCTATCTCAGCGCTTATTTATTTAC 573
QY 405 -----TCGAGGCGGACACCGCGGCGCAAGCTCGTTCCTTATATCTCATTTAGATG 459
    || || || || || || || || || || || || || || || || || || || || ||
Db 574 AGTACTGTGGCTACGCTTCCAACTCTGGCTCTCTTTATTAATTTGCATCCAAATG 633
QY 460 ATCTCCGAGCGCGGAGATTCAATCCCATCTTTTGGAGGCTGCGCAATTAACAGC 519
    || || || || || || || || || || || || || || || || || || || || ||
Db 634 ATTTCAGACGACGACAAATTTCCATATATTTGAGGAGAAATGCGACGACAAATTAAGTAC 653
QY 520 GGGAGTCAATTTCTTCCGACATGTACATGCTGAGCTGGAGAGACTAGTTGGGCGCAAA 579
    || || || || || || || || || || || || || || || || || || || || ||
Db 694 AACCGAGATCTGACCACTTCTAGCGTAAATTAACACTTGAGAAATAGTTGGGAGACTT 753
QY 580 TCCACGCAATCCGACGATGTACGATGCGGCTTTTAAATTAACCAATTTGCGTTGGTATA 639
    || || || || || || || || || || || || || || || || || || || || ||
Db 754 TCCACTGCAATTCAGAGCTTAACCAAGAGCGCTTGTGCTAGTCCAAATTCACGCAAGA 813
QY 640 TCCACCGTAACCTCTGTGACGTTGAGCAATGTTGCGACGCTGATCGGACCTTAGCGATG 699
    || || || || || || || || || || || || || || || || || || || || ||
Db 814 CGTAATGTTCCAAATTCAGTGTGTAGATGTGATATTAATCCCTATCATATAGCTCTC 873
QY 700 ATGTGTTTGTATGTAGGAGACGACATCTTC 731
    || || || || || || || || || || || || || || || || || || || || ||
Db 874 ATGGTGTATAGATGCGCACCTCCACCATCTC 905

```

```

RESULT 10
US-09-147-208-47
; Sequence 47, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 47:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1855 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-47

```

```

Query Match          9.2%; Score 77.2; DB 4; Length 1855;
Best Local Similarity 49.1%; Pred. No. 5,8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

```

```

QY 124 CAAGATTGTGTGTGGTGAACATCAACATTCAGGGGGAGAGCTGATACGCGCCCATTC 183
    || || || || || || || || || || || || || || || || || || || || ||
Db 274 CAACGGTTATTTTATTTAATTAATCTCAATTCATGCAAGACTTCTGTTACATTAACGCTG 333
QY 184 GACGTTACCAATCTGTACGTGTGGCTTACCAAGAGCGGACCAATCTCACTTTTGGC- 242
    || || || || || || || || || || || || || || || || || || || || ||
Db 334 GATGTCAACCAATGCAATATGTGTGGCTACCGCTACCGTGGAAATAGGCAATATTTCTTCAT 393
QY 243 -----CGACGACCAAGCGCGCGGAAAGCATCTTTCACCGGACCAACCA 291
    || || || || || || || || || || || || || || || || || || || || ||
Db 394 CCTGACATCAGGAAGATGCAAGCAATCACTCATCTTTTCACTGATGTTCAAAATGCA 453
QY 292 TCCTCTCTCCCATTCACCGAAGCTACACAGATCTGAGACGATAGCGCGCTCAT--AGG 348
    || || || || || || || || || || || || || || || || || || || || ||
Db 454 TATCATTTGCTTTGGTGTGTAATTAATGATAGACTTTGAACAATCTGCTGTGTAATCTGAGA 513
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCAATCCGTCGTGCGGCTTCGTTA--- 404
    || || || || || || || || || || || || || || || || || || || || ||
Db 514 GAAATATGAGTTGGGAAATGTGTCACATAGAGAGGCTATCTCAGCGCTTATTAATTAC 573
QY 405 -----TCGAGGCGGACACCGCGGCGCAAGCTCGTTCCTTATATCTCATTTAGATG 459
    || || || || || || || || || || || || || || || || || || || || ||
Db 574 AGTACTGTGGCTACGCTTCCAACTCTGGCTCTCTTTTATTAATTTGCAATTCGAATG 633
QY 580 TCCACGAGCGCGGAGATTCAATCCCATCTTTTGGAGGCTGCGCAATTAACAGC 519
    || || || || || || || || || || || || || || || || || || || || ||
Db 634 ATTTCAGAGCAGCAAGATTCCAATATATTTGAGGAGAAATGCGACGAAATTAAGTAC 693
QY 520 GGGAGTCAATTTCTTCCGACATGTACATGCTGAGCTGGAGAGACTAGTTGGGCGCAAA 579
    || || || || || || || || || || || || || || || || || || || || ||
Db 694 AACCGAGATCTGACCAAGATCTTACGTAATTAACACTTGAGAAATAGTTGGGAGACTT 753
QY 580 TCCACGCAATCCGACGATGTACGATGCGGCTTTTAAATTAACCAATTTGCGTTGGTATA 639
    || || || || || || || || || || || || || || || || || || || || ||
Db 754 TCCACTGCAATTCAGAGCTTAACCAAGAGCGCTTGTGCTAGTCCAAATTCACCTGCAAGA 813
QY 640 TCCACCGTAACCTCTGTGACGTTGAGCAATGTTGCGACGCTGATCGGACCTTAGCGATG 699
    || || || || || || || || || || || || || || || || || || || || ||
Db 814 CGTAATGTTCCAAATTCAGTGTGTAGCATGTGATATTAATCCCTATCATATAGCTCTC 873
QY 700 ATGTGTTTGTATGTAGGAGACGACATCTTC 731
    || || || || || || || || || || || || || || || || || || || || ||
Db 874 ATGGTGTATAGATGCGCACCTCCACCATCTC 905

```

```

RESULT 11
US-09-147-208-54
; Sequence 54, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/147,208
 FILING DATE: 02-MAR-1999
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Rudolph, John R.
 REGISTRATION NUMBER: 38,003
 REFERENCE/DOCKET NUMBER: 7841-76
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 364-7311
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1855 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-147-208-54

Query Match 9.2%; Score 77.2; DB 4; Length 1855;
 Best Local Similarity 49.1%; Pred. No. 5.8e-16;
 Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

124 CAAGATTGTTGTTGGTGAAGTCAACCAATCAGGGGGAGACTGATACGGCCGATC 183
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 274 CAACGGTTTATTTAGTTGAAGTCAACCAATCAGAGAGCTTCTGTTACATTAAGCGGTG 333
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 184 GACGTACCAATCTGTAGTGGTGGTCAACGAGGAGCCATCTTACTTTTGGG- 242
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 334 GATGTCACCAATGCATATGTGTGCTGACCTGCTGGAATAGGCGCATTTCTTTCAT 393
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 243 -----CGACGACCAAGACGGCGGGAAGGCACTCTTCCACCGGACCAACGAGA 291
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 394 CCGACCAATCAGAGAGATGAGAACCAATCACTCTTTTCCATGATGTTCAAAATGCA 453
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 292 TCCCTCTCCCATTCACCGGGAAGTACACAGATCTGGAGCGATACGGCGGAT---AGG 348
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 454 TATTAATTCGCTTTGGTGTATATATGATAGACTTGAACAACCTTGCTGTAATCTGAGA 513
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 349 GACGATCCCTCTGGGTATAGAGAACTCAATCCGCTCGGCGCTTCGTTA----- 404
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 514 GAAATATCGAATGTGGAAATGTCTACTAGAGAGGCTATCTCGCGCTTATTTATAC 573
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 405 -----TCAGGCGGACGACACCGGGCCCAAGCTGCTCCCTTATATATCTCATTCAGATG 459
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 574 ACTACTGGTGGACACCTCCACTCTGCTGCTCTCTTATATTTGATCCCAATG 633
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 460 ATCTCGAGGGCGGAGATTCATCCCATCTTTTGGAGGGGCGCCATATTAACAGC 519
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 634 ATTTAGAGACGACGAGATTCATATTTAGAGGAGAAATGCGACAGATTAAGTAC 693
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 520 GGGGAGATTTCTTCGACATGTACATGCTGAGCTGAGAGTATGGGGCCACAA 579
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 694 AACCGAGATCTGACACAGATCTGAGCAATTAACCTTGAGAAATATGGGGGAGATT 753
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 580 TCCAGGCAAGTCCAGCAGTCTAGCGATGCGCTTTTAATTAACCATTTCCGTTGGGTATA 639
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 754 TCCATGCAATTCAGAGCTTAACCAAGAGGCTTGTAGTCCCAATTCACATTCGAAAGA 813
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 640 TCCACCGGTAATCTGTGAGCTTGAGCAATGTCGCAAGGATGAGCAGGTTAGCGATC 699
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 814 CGTAATGCTCCAAATTCAGTGTGACGTGAGATATATTAATCCATATAGCTCTC 873
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 700 ATGTTGTTGTATGAGGACCGACCATTTTC 731
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 874 ATGCTATAGATGGCGCACTCCACCATCTGC 905
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 12
 US-09-147-208-23
 Sequence 23; Application US/09147208
 Patent No. 633303
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Antiviral Ricin-Like Proteins
 NUMBER OF SEQUENCES: 71
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BERSKIN & PARR
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/147,208
 FILING DATE: 02-MAR-1999
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Rudolph, John R.
 REGISTRATION NUMBER: 38,003
 REFERENCE/DOCKET NUMBER: 7841-76
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 364-7311
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1879 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-147-208-23

Query Match 9.2%; Score 77.2; DB 4; Length 1879;
 Best Local Similarity 49.1%; Pred. No. 5.8e-16;
 Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

124 CAAGATTGTTGTTGGTGAAGTCAACCAATCAGGGGGAGAGACTGATACGGCCGATC 183
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 298 CAACGGTTTATTTAGTTGAAGTCAACCAATCAGAGAGCTTCTGTACATTAAGCGGTG 357
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 184 GACGTTCAATCTGTAAGTGTGCTTACCAAGCAGGCGACCAATCTACTTTTGGG- 242
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 358 GATGTCACCAATGCATATGTGTGCTGCTACCGCTGGAATAGCGCATATTTCTTTCAT 417
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 243 -----CGACGACCAAGACGGGGCGGAAGGCAATCTCTTCCACCGGACCAACCA 291
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 418 CCGACCAATCAGAGAGATGAGAAACCAATCACTCATCTTTTCCATGATGTTCAAAATGCA 477
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 292 TCCCTCTCCCATTCACCGGGAAGTACACAGATCTGAGAGGATACCGCGGAT---AGG 348
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 478 TATTAATTCGCTTTGGTGTATATATGATAGACTTGAACAACCTCTGCTGTAATCGAGA 537
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 349 GACGATCCCTCTGGGTATAGAGAACTCAATCCGCTCTCGGCGCTTGTGTTA----- 404
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 538 GAAATATTCAGATGTGGAAATGTGTCACTAGAGAGGCTATTCACGCTTATTAATTTAC 597
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 405 -----TCAGGCGGACGACACCGGGCCCAAGCTGCTTCCCTTATATCTCATTTAGATG 459
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 598 AGTACTGTGGACACTAGCTTCCACTCTGGCTGCTTCTTATTAATTTGATCAATG 657
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 460 ATCTCGAGGCGGAGATTCATCCCATCTTTTGGAGGGGCGCCATTAATTAACAGC 519
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 658 ATTTCAAGACGAGAAATTCATATATGAGAGGAAATGCGACAGAAATATAGGTAC 717
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 520 GGGGATCATTTCTTCCGCAATGATACATGCTGAGCTGAGACTGATGTTGGGGCCACAA 579
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

Db 718 AACCGAGATCTGCACAGATCTAGCGTATTAACCTTGAGATAGTGGGAGACTT 777
QY 580 TCCACGACAGTCCAGACGTCTACGAGATGGCTTTTAAATACCATTTGGTGGTATA 639
Db 778 TCCACTGCAATTCAGATGTAAACCAAGGACCTTTCTAGTCCAAATTCACCAAGA 837
QY 640 TCCACCGTAACTCTGAGCTTGTAGCAATGTTCGACGAGTATGCGAGCTTAGCATC 699
Db 838 CGTATGCTTCCAAATTCAGTGTGTAGATGTGATATTAATCCATCATAGCTCTC 897
QY 700 ATGTTGTTGATGTAGGAGCAGCATCTTC 731
Db 898 ATGGTGTATAGATCGCACCTCCACCATCTGC 929

RESULT 13
US-09-147-208-24
; Sequence 24, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Rictin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147, 208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-24

Query Match 9.2%; Score 77.2; DB 4; Length 1879;
Best Local Similarity 49.1%; Pred. No. 5.8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

QY 124 CAAGATTGTGTGGTGAACCTCACCAATCAGGGGAGAGACTGCATCAGCGCCATC 183
Db 298 CAACGGTTATTAGTTGAACCTCTCAATATCATGCAAGCTTCTGTATCATTAAGCCTG 357
QY 184 GACCTTACCAATCTGTACGTGGTGGCTTACCAAGCAGCACAATCTCTTTTGGC- 242
Db 358 GATGTACCAATGCATGTGTGGTGGCTTACCGCTGAATAGCCGATATTTCTTCAT 417
QY 243 -----CGACGACACAGACGGCGGAAAGCATCTCTACCGGACACACAGA 291
Db 418 CCTGACATCAGGAAGATGCAAGAACCAATCATCATCTTTTCACTGATGTTCAAATGCA 477
QY 292 TCTCTCTCCCATTCACCGGAAGCTACACAGATCTGGAGGAGATACGCCGCTCAT--AGG 348

Db 478 TATACATTCGCCCTTTGGTGGTATTTATGATAGACTTGAAACAACTGCTGTAATCTGANA 537
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCAATTCATCCGCTCGGGCGCTTGTA---- 404
Db 538 GAAATATGAGTGGGAATGGTCCACTAGAGAGAGGTATCTCAGCGCTTATATATTAC 597
QY 405 -----TCCAGGGCGGACAGACCCGGGCCAAAGCTCGTCCCTATATCTCATTTGAGATG 459
Db 598 AGTACTGGTGGCACTAGCTTCCAACTCTGGCTCGTCTTATATATTTGCAATCCAAATG 657
QY 460 ATCTCCGAGCCCGAGATTCATCCCATCTTTGGAGGGCTCGCCAAATACATTAACAGC 519
Db 658 ATTTCAAGAGCAGCAGATTTCAATATATTAGAGGAGAAATCGCAGCAGATTTAGTTC 717
QY 520 GGGAGCTATTTCTCCGACATGTACATGCTCGACCTGAGACATAGTTGGGCCAACAA 579
Db 718 AACCGAGATCTGCACCAATCTAGCTTAATTAACATTCAGAAATAGTTGGGGAGACTT 777
QY 580 TCCACGACAGTCCAGAGTCTACGATGGCGTTTAAATACCATTTGGTGGTATA 639
Db 778 TCCACTGCAATTCAGATGTAAACCAAGAGCCTTTGCTAGTCAATTCACCTGCAAGA 837
QY 640 TCCACCGTAACTCTGAGCTTGTAGCAATGTTCGCGAGCTGATCGCCAGCTTAGCATC 699
Db 838 CGTATGCTTCCAAATTCAGTGTGTAGCATGTGAGATATTAATCCATCATAGCTCTC 897
QY 700 ATGTTGTTGATGTAGGAGCAGCATCTTC 731
Db 898 ATGGTGTATAGATCGCACCTCCACCATCTGC 929

RESULT 14
US-09-147-208-25
; Sequence 25, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Rictin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147, 208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-25

Query Match 9.2%; Score 77.2; DB 4; Length 1879;

Best Local Similarity 49.1%; Pred. No. 5,8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

```

QY 124 CAAAGATTGTGTGGTGAACACACCATCAGGGGGAGAGCTCCATCAGCGCCGATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 CAAGGGTTATTTTATGTAAGTCTCAATCATGAGAGCTTCTTCTTATTAAGCGG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 GAGCTTACCAATCTAGCTAGCTGAGCTTACCAAGAGGAGACCATCTCACTTTTGGC- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 GATGTACCAATCATATGAGTGGCTACCGCTGCGAATAGCCCATATTTCTTCAAT 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 -----CGAGCACCAAGAGCGCGGAGAAAGCATCTCTTCAACCGGACCAACCA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 CCTGACATTCAGGAGATGAGAGCAATCATCTCTTTCATCTGATTTCAAAATGCA 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 TCTCTCTCCATTTACCGGAGATCAGACATCTGAGAGGATAGCGCGTAT---AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 TATACATTCGCTTTTGTGTGATTTATGATAGCTTGAACACTTCTGTGATATTCAG 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTA---- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 GAAATATGAGTTGGGAAATGCTCCACTAGAGAGGCTATCTCAGCGCTTATATTTAC 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 -----TCAGGCGGAGACCGCGGCCAAGCTCGTCCCTTAATAATCTCATTCAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598 AGTACGTGGGCACTGAGCTTCACACTGCGCTGCTCTTATATAATTTGACCAAAATG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ATCTCCGAGGCGCGAGATTCATTCATCTTTTGGAGGGGCTCGCCATATCATTTACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 658 ATTTCAGAGACGACAGATTTCCATATATTAGAGGAGAAATGCGCAGCAATTTAGTAC 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 GGGAGCATTTCTTCCCATGATGATGCTGAGCTGAGAGCTGAGATTGGGGCCAAACA 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 AACCGAGATCTCAGCAATCTCTAGCTAATTTACATTTAGAAATTTGGGGAGACATT 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TCCAGCAGTCCAGAGCTTACGATGAGTGGCGTTTAAATACCATTTGGTGGGTATA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 TCCAGTGCATTTCAAGTCTTACCAAGAGCTTTGCTGATTCATTCATTCAGCAAAA 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 TCCAGCAGTCTTCTGAGCTTGAAGCAATGTCGAGCTGAGTCCAGCTTACGAGTAC 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 CGTAATGTTCCAAATTCAGTGTGATGATGATATTAATCCCTATCATAGCTCTC 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 ATGTTGTTGATGATGAGGAGCAGCATCTTC 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 ATGTTGATAGATGCGCACCCTCCACCATCTGC 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

US-08-425-336-11
Sequence 11, Application US/08425336

Patent No. 5621083

GENERAL INFORMATION:

APPLICANT: Better, Marc D.

APPLICANT: Carroll, Stephen F.

APPLICANT: Studnika, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower, 233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,336

```

? FILING DATE: 18-APR-1995
? CLASSIFICATION: 530
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/064,691
? FILING DATE: 12-MAY-1993
? APPLICATION NUMBER: US 07/901,707
? FILING DATE: 19-JUN-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/787,567
? FILING DATE: 04-NOV-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Meyers, Thomas C.
? REGISTRATION NUMBER: P-36,989
? REFERENCE/DOCKET NUMBER: 31394
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 312/474-6300
? TELEFAX: 312/474-0448
? TELEX: 25-3856
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 813 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? US-08-425-336-11

```

Query Match 7.9%; Score 66.4; DB 1; Length 813;

Best Local Similarity 48.2%; Pred. No. 1.7e-12; Matches 292; Conservative 0; Mismatches 296; Indels 18; Gaps 3;

```

QY 128 GATTGTGCTGTGGTGAACATCAATCAGAGGGGAGAGCTGATCAGCGCCGATCAGC 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 GTTGTGTTGGTGAACGCTTTCAATGACATGACAGTTGGCGGAATAGCTATAGATG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 188 TTACCAATCTGATGCTGCTTACCAAGCAGCGCAATCTTACTTTTGGCGAGC 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 TTACCAAGTTATGTTGCTGCTATCAAGTAAAGAAACAGATCTTACTTTTAAAGATG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 248 CACGAGCGGGCGGAAAGGATCTCTACCGCGCACACAGATCTCTCCCATTC 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 CTCGAGTCTGCTTACCAAGCGCTCTTCAAAACACATTTAAACAGACTTCATTTTG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 CCGAAGCTTACAGATCTGAGCGAATAGCGGCTCATAGGACAGCATCCCTGCGGTA 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GCGGAGCTATCCCTCGCTGGAAGTGAAGCATATAGAGACAAACAGACTTGGCA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 368 TAGAGACCTATTCATCCCTCTCGCGCTTCTTATCCAGGGCGAGCA----- 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 TTGAACCATTAAGGATTTGGCATCAAGAACTTGATGAATAATGCATAGACAAATTATAAC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 CCGGCGCCAGAGCTGTTCCCTTAAATCCATTCAGATGATCTCGAGGCGCGAGAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 CACGAGATAGCTAGTCTCTATTTGTTATTTCAAAATGCTGCGAAGCAGCTGAT 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 TCAATCCATCTTTTGGAGGCTCGCAATATCAACAGCGGGAGCTATTTCTCCG 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 T-----CACCTTATGAGAACCAATTAAGAAATCTTCAACAGAGATTCGCCGG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 539 ACATGTATACCTGAGCTGAGACTAGTTGGGCCAACAATCCAGCAATCCAGACT 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 CGAATTAATCAATAGCTTGAAGATAAATGGGTAAATCTGCTTCAAGATCCGAGCAT 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 599 C---TACGATGGGCTTTTAAATTAACCATTTGCTGGGTATATCCAGCGTAACTTCG 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 623 CAGGTGCAATGCAATGTTTTCGAGGAGCTTGAATTTGGAACGTGCAATATGGCAAAAT 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 656 TGACGTTGAGCAATGTTGCGAGCTGATCGCAGCTTAGCATCATGTTGTTGATGTA 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 ACTATGTCAGCGAGTGTATCAATTAACCCAAATATAGACACCTTGAATTCGTGATA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 716 GGGAGC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Thu Mar 27 07:06:57 2003

us-09-627-165e-15.rni

Page 11

Db 743 AAGATC 748

Search completed: March 26, 2003, 19:43:05
Job time : 58.4724 secs

•
•
•
•

OY 181 ATCGAGCTTACCAATCTGACCTGGTGTATACGACGAGCGACCAATCCTACTTTTG 240
 Db 208 ATCGAGCTTACCAATCTGACCTGGTGTATACGACGAGCGACCAATCCTACTTTTG 267
 OY 241 CGCGACGACCAAGAGCGCGCGGAAAGGACATCTCTTCACCGGACACACAGATCCTCTTC 300
 Db 268 CGCGACGACCAAGAGCGCGCGGAAAGGACATCTCTTCACCGGACACACAGATCCTCTTC 327
 OY 301 CGATTCACCGGAAGTATACACAGATCTGGAGGATACGCGGTCATATAGGACAGATCCTT 360
 Db 328 CCAATTCACGGAAGTATACCTGATCTGGAGGATACGCGGACATATAGGACAGATCCTT 387
 OY 361 CTGGGTATAGAGACATCATTCATTCCTCTCGGCGCTTGCTATATCCAGGCGGACGACC 420
 Db 388 CTGGGTATAGAGACATCATTCATTCCTCTCGGCGCTTGCTATATCCAGGCGGACGACC 447
 OY 421 CGGCGCCAAAGCTGCTCCCTTATATCTCATTCAGATGATCTCGAGCGCGGAGATTC 480
 Db 448 CCGTACCAAGCTGCTCCCTTATATCTCATTCAGATGATCTCGAGCGCGGAGATTC 507
 OY 481 AATCCCATCTTTGGAGGGCTGCGCAATCATTAACGAGCGGGAGTCAATTTCTCCGAC 540
 Db 508 AATCCCATCTTATGGAGGGCTGCGCAATCATTAACGAGGGGCTCAATTTCTCCGAC 567
 OY 541 ATGTACATGCTCGAGCTGGAGCTAGTTGGGGCCAAATTCACGACGACGACGACTCT 600
 Db 568 GTGTACATGCTGGAGCTGGAGCGAGTTGGGGCCAAATTCACGACGACGACGACTCT 627
 OY 601 ACGGATGGCGTTTATATACCCATTTCCGTTGGGTATATCCACGCGTAATCTCGTAGC 660
 Db 628 ACCGATGGCGTTTATATACCCAAATTCGTTGGGTATATCCCGGTAATCTCGTAGC 687
 OY 661 TTGACCAATGTTTCGGAGCTGATGCGCGAGCTTAGCGATGTTGTTATGAGGAGC 720
 Db 688 TTGACCAATGTTTCGGAGCTGATGCGCGAGCTTAGCGATGTTGTTATGAGGAGC 747
 OY 721 CGACCATCT 729
 Db 748 CGGCGCATCT 756
 RESULT 2
 US-09-347-064-1
 ; Sequence 1, Application US/09347064A
 ; Patent No. US20020045208A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eck, Jurgen
 ; APPLICANT: Schmidt, Arno
 ; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
 ; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
 ; FILE REFERENCE: 09282-5
 ; CURRENT APPLICATION NUMBER: US/09/347,064A
 ; CURRENT FILING DATE: 1999-07-02
 ; EARLIER APPLICATION NUMBER: PCT/EP98/00009
 ; EARLIER FILING DATE: 1998-01-02
 ; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 762
 ; TYPE: DNA
 ; ORGANISM: Viscum album
 ; US-09-347-064-1

Query Match	76.5%	Score 640.2	DB 10	Length 762
Best Local Similarity	92.7%	Pred. No. 2.7e-216		
Matches 672	Conservative 0	Mismatches 53	Indels 0	Gaps 0
QY	1	CATCAGACGAGCGGAGCAATTTCCGGTTCATCACGCTTCTCGGAATATTGCTCTA	60	

Db	34	CACCAACCAACCGGTGAGAAATATTTCCGGTTACACGCTTCCGAGATTATCTTCA	93
QY	61	AGCGGAAGCTTTTCCATAGATATACCACTCTTGGCTAGTCTAGCATCCCGCTCGGAT	120
Db	94	AGCGGAAGCTTTTCCATAGATATACCACTCTTGGCTAGTCTAGCATCCCGCTCGGAT	15
QY	121	GCGCAAGATTTGTGTGTGTAAGTCAACCAATGAGGGGAGACTCGATCAGGCGCC	180
Db	154	GCGCAAGATTTGTGTGTGTAAGTCAACCAAGGGGAGACTCGATCAGGCGCC	21
QY	181	ATGAGGTTACCAATCTGTAGCTGGTGGCTTACCAACGAGGCAATCCACTTTTG	240
Db	214	ATGAGGTTACCAATCTGTAGCTGGTGGCTTACCAACGAGGCAATCCACTTTTG	27
QY	241	CGGAGGCAACAGAGCGGCGGAAAGGATCTCTACCGGCAACCAAGATCTCTCTC	300
Db	274	CGGAGGCAACAGAGCGGCGGAAAGGATCTCTACCGGCAACCAAGATCTCTCTC	33
QY	301	CCATTACCGGAAGCTACACAGATCTGGAGCGATGCGCGGTCATAGGAGCCAGATCCCT	360
Db	334	CCATTACCGGAAGCTACACAGATCTGGAGCGATGCGCGGATAGGAGCCAGATCCCT	39
QY	361	CTGGGTAATGAGAACTCATTCATCCGCTCTCGGCGCTTCGTTATCCAGGGGCAAGCAC	420
Db	394	CTGGGTAATGAGAACTCATTCATCCGCTCTCGGCGCTTCGTTATCCAGGGGCAAGCAC	45
QY	421	CGGAGCCAGAGCTGTTCCCTTAATATCCCTCATTCAGATGATCTCGAGGCGCGAGATTG	480
Db	454	CGAGCCAGAGCTGTTCCCTTAATATCCCTCATTCAGATGATCTCGAGGCGCGAGATTG	513
QY	481	AATCCCATCTTTTGGAGGGCTCGCCATATACATTAAAGACGCGGGAGTCAATTTCTCCGAC	540
Db	514	AATCCCATCTTTTGGAGGGCTCGCCATATACATTAAAGAGGGGCGTCATTTCTCCGAGAC	573
QY	541	ATGTACATCTGTGAGCTGGAGACTAGTAGTTGGGGGCAACAAATCCAGCAGATCCAGAGCT	600
Db	574	ATGTACATCTGTGAGCTGGAGACTAGTAGTTGGGGGCAACAAATCCAGCAGATCCAGATTC	633
QY	601	ACGAGTGGGCTTTTAATTAACCAATTCCTGGTATATCCACCGGTAACCTCTGTCAGC	660
Db	634	ACGAGTGGGCTTTTAATTAACCAATTCCTGGTATATCCACCGGTAACCTCTGTCAGC	693
QY	661	TTGAGCAATGTTCCGAGCGTATGCGCAGCTTAGCGATCATGTTGTTGATAGTAGGAG	720
Db	694	TTGAGCAATGTTCCGAGCGTATGCGCAGCTTAGCGATCATGTTGTTGATAGTAGGAG	753
QY	721	CGACC 725	
Db	754	CGCCC 758	

```

RESULT 3
US-09-785-921A-1
? Sequence 1, Application US/09785921A
? Patent No. US2002009434A1
? GENERAL INFORMATION:
? APPLICANT: Keener, William K.
? APPLICANT: Ward, Thomas E.
? TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
? FILE REFERENCE: LIT-PI-529
? CURRENT APPLICATION NUMBER: US/09/785,921A
? CURRENT FILING DATE: 2001-02-15
? NUMBER OF SEQ ID NOS: 17
? SEQ ID NO 1
? LENGTH: 1698
? TYPE: DNA
? ORGANISM: Rictinus communis
US-09-785-921A-1

Query Match          9.4%; Score 78.8; DB 10; Length 1698;
Best Local Similarity 49.2%; Pred. No. 2.7e-17;
Matches 311; Conservative 0; Mismatches 297;
Indels 24; Gaps 3;

```



```

Db      388 CCTGACATTCAGGAGATGAGCAGAACATTCATCTTTTCACATGATGTTCAAAATGCA 447
QY      232 TCCTCTCCCTCCATTCACCGAAGCTACACAGATCTGGAGCATACCGCGTCAAT---AGG 348
Db      448 TATACATTCGCTTTGGTGGTAAATATATATAGACTGAAACAACTGGTGGTAAATCTGAGA 507
QY      349 GACGAGATCCCTCTGGGTATAGAGAACTCAATTCAGTCCGCTGGGCTTGTATTA---404
Db      508 GAAATATGAGATTGGGAAATGATGCTCAGAGAGGCTATCTCAGCGCTTTATTTATAC 567
QY      405 -----TCCAGGCGGAGCAGACCCGCGCCCAAGCTGTTCCCTTAATATCCATTCAGATG 459
Db      568 AGTACTGGGCACTCAGACTTCACACTCTGCGTCTGCTTTATTAATTTGATCCAAATG 627
QY      460 ATCTCCGAGGCGCGAGATTCATTCCTTTTGGAGGCGTCCGCAATATTAACACG 519
Db      628 ATTTGAGAGCAGCAAGATTCCTCAATATTTGAGGAGAAATGCGCAGAGAAATAGGTAC 687
QY      520 GGGAGTCAATTTCTCCGACATGATGCTCGAGCTGAGAGACTAGTGGGCGCAACA 579
Db      688 AACCGAGATCTGCACAGATCTAGCGTAAATTAATGATGAGATTTGGGGAGACTT 747
QY      580 TCCACGCAAGTCCAGAGCTACAGGATGCGCTTTTAAATTAACCAATTCGTTGGGTATA 639
Db      748 TCACCTGCAATTCAGAGATCTAACCAAGAGCGCTTGTCTAGTCCAAATTCACATGCAAGA 807
QY      640 TCCACCGGTAACTTGTGATGCTGAGCAATGTCCGCGAGTATGCGCAGCTTAGCCATC 699
Db      808 CGTAAATGTTCCAAATTCAGTGTGATGATGATATATATATCCATATCATAGCTCTC 867
QY      700 ATGTTGTTGTATGTAGGAGCAGCATCTTC 731
Db      868 ATGGTATATAGATGGCAGCTCCACCATCTGC 899

```

RESULT 6

```

US-09-785-921A-10
; Sequence 10, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 10
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: R1c1nus communis
US-09-785-921A-10

```

```

Query Match 9.0%, Score 75.6; DB 10; Length 1805;
Best Local Similarity 48.9%; Pred. No. 3,7e-16;
Matches 309; Conservative 0; Mismatches 299; Indels 24; Gaps 3;

```

```

QY      124 CAAGATTTGTGTGGTGAATCACCATCAGGGGAGAGACTGCATACGCGCCATC 183
Db      276 CAAGGTTTATTTATGTTAGTCAATCAATCATGAGAGCTTTGTGTACATTAAGCGGTG 335
QY      184 GACGTTACCAATCTGTAGTGTGCTTACCAAGAGCGGAGCAATCTTACTTTTGGG- 242
Db      336 GATGTACCAATGATATGTGTGCTTACCGTCTGTAATATAGGCAATTTCTTTTAT 395
QY      243 -----CGACGACAGCAGCGCGCGGAGAAAGCATCTTTACCGGAGCAACACAGA 291
Db      396 CCGTCAATCAGAGATGAGAGAGCAAGCAATCACTCTTTTACATGATGTTCAAAATGCA 455
QY      292 TCCTCTCTCCATTCACCGGAGATGACAGATCTGGAGCGATACGCGGCTCAT---AGG 348

```

```

Db      456 TATACATTCGCTTTGGTGGTAAATATATAGACTTGAACAACTTCTGTTATCTGAGA 515
QY      349 GACGAGTCCCTCTGGGTATAGAGAACTCATTCATCCGCTCGCGCGTGGTAA-----404
Db      516 GAAATATGAGTTGGGAAATGCTCCACTAGAGAGGCTATCTCAGCGCTTATTTATTTAC 575
QY      405 -----TCCAGGCGGAGCAGACCCGCGCCCAAGCTGTTCCCTTAATATCCATTCAGATG 459
Db      576 AGTACTGTGTGCACTCAGCTTCCACTCTGCTGCTCTTTATATATTTGATCCAAATG 635
QY      460 ATCTCCGAGGCGCGAGATTCATTCCTTTTGGAGGCGTCCGCAATATTAACACG 519
Db      636 ATTTCAAGACAGCAAGATTCCTCAATATATAGAGGAGAAATGCGCAGAGAAATAGGTAC 695
QY      520 GGGAGTCAATTTCTCCGACATGATGCTCGAGCTGAGAGACTAGTGGGCGCAACA 579
Db      696 AACCGAGATCTGCACAGATCTAGCGTAAATTAATGATGAGATTTGGGGAGACTT 755
QY      580 TCCACGCAAGTCCAGAGCTTACAGATGCGCTTTTAAATTAACCAATTCGTTGGGTATA 639
Db      756 TCACCTGCAATTCAGAGTCTAACCAAGAGCGCTTGTCTAGTCCAAATTCACATGCAAGA 815
QY      640 TCCACCGGTAACTTGTGATGCTGAGCAATGTTCGCGAGCTATGCGCAGCTTAGCCATC 699
Db      816 CGTAAATGTTCCAAATTCAGTGTGATGATGATATATATATCCATATCATAGCTCTC 875
QY      700 ATGTTGTTGTATGTAGGAGCAGCATCTTC 731
Db      876 ATGGTATATAGATGGCAGCTCCACCATCTGC 907

```

RESULT 7

```

US-09-765-527-246
; Sequence 246, Application US/09765527
; Patent No. US2002000638A1
; GENERAL INFORMATION:
; APPLICANT: Belter, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; Fusion Proteins and BPI-Derived Peptides
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

```

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

```

INFORMATION FOR SEQ ID NO: 246:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 813 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```



```

;      FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "gelonin"
;      SEQUENCE DESCRIPTION:  SEQ ID NO: 246:
US-09-765-527-246

```

Query Match	7.98;	Score 66.4;	DB 10;	Length 813;
Best Local Similarity	48.28;	Pred. No. 4.6e-13;		
Matches 292; Conservative	0;	Mismatches 296;	Indels 18;	Gaps 3;

OY	128	GATTGGTGGTGGAGCTACCAATCAGGGGGAGACTCGATCAAGGCCCGCATCGACG	187
Db	149	GTTCGTTTTGGTAGCCCTTTCAAAATGACAAATGAGACGTTGGCGGAAATACCTATTAGATG	208
OY	188	TTACCAATCTGTACGTGGTGGCTTTACCAAGCAGCGGCCACCATCCATCTTTTGGCGACG	247
Db	209	TTACAAAGCTTTATGTGTGTGGCTATCAAGTAGAAACAGATCTTACTTCTTTAAAGATG	268
OY	248	CACCAGACGGCGGGAAAGGCATCTCTTCAGCGGCACGCCACGATCCTCTCCCATTTCA	307
Db	269	CTCCAGATGCTGCTTACGAAAGGCCCTCTTCAAAAACACATTTAAACAGACTTCATTTTG	328
OY	308	CCGGAACCTACACAGATCTGTGAGCCATACGCCGGTCATATGGAGCACAGATCCCTGTGGGTA	367
Db	329	GGCGACACTATCCCTCGCTGGGAAGGTGAAGAGGCATATAGAGAGACAAAGACTTTGGGCA	388
OY	368	TAGAGAGAACATTCATCAATCCGTCCTGGCGCTCGTTATCCAGGGCGGCAGCA-----	418
Db	389	TTGAAACATTAAAGGATTTGGCATTAAGAAACTTGATGAAATGCGATTAGACAAATTATAAC	448
OY	419	CCCGGGCCCAAGCTCGTTCCCTTATATATCCATTCAGATGATCCGAGGCCGAGAT	478
Db	449	CAACGGAGATAGCTAGTCTCTATTTGGTTGTTATTCAATGTTGTCTGAAACACTTGAT	508
OY	479	TCATCCCATCTTTGGAGGGCGTCGCCAATACATTAAACGGGGGAGTCATTTCTCCG	538
Db	509	T-----CACCTTTATTTAGAAACCAATATGAAATACCTTTCAACAGAGAAATGGCCGG	562
OY	539	ACATGTACATGCTCGAGCTGAGACTAGTATGGGGCCCAACATCCACGAGTCCAGCAGT	598
Db	563	CGAATATATCAATCAGACCTTTGAGAAATTAATGGGGTAAACTCTCGTTCCAGATCCGGACAT	622
OY	599	C---TACGATGGCGTTTTTAAATTAACCATTTTCGGTGGGTATATGCCACGGTAACTTCG	655
Db	623	CAGGTGCAATGGAATGTTTTCGGAGGCGATTGAATTTGGAACGTGCAAAATGGCAAAAAT	682
OY	656	TGACGTTGACCAATGTTCCGCGACGTGATGCCAGCTTAGCGCATGATGTTGTTGATGTA	715
Db	683	ACTATGTACACGCGAGTTGATCAAGTAATAACCCAAATATGACACTCTTGAAGTGGTGATTA	742
OY	716	GGGAGC 721	
Db	743	AAGATC 748	
<p>RESULT 8</p> <p>US-09-765-527-258</p> <p>Sequence 258, Application US/09765527</p> <p>Patent No. US2002000638A1</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Belter, Marc D.</p> <p>TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides</p> <p>NUMBER OF SEQUENCES: 265</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borzun</p> <p>STREET: 6300 Sears Tower, 233 South Wacker Drive</p> <p>CITY: Chicago</p> <p>STATE: Illinois</p> <p>COUNTRY: United States of America</p> <p>ZIP: 60606-6402</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p>			

```

1  COMPUTER: IBM PC compatible
2  OPERATING SYSTEM: PC-DOS/MS-DOS
3  SOFTWARE: Patentia Release #1.0, Version #1.25
4
5  CURRENT APPLICATION DATA:
6  APPLICATION NUMBER: US/09/765,527
7  FILING DATE: 18-Jan-2001
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: 08/621,803
10 FILING DATE: <unknown>
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Borzun, Michael F.
13 REGISTRATION NUMBER: 25,447
14 REFERENCE/DOCKET NUMBER: 27129/33199
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 312/474-6300
17 TELEFAX: 312/474-0448
18
19 INFORMATION FOR SEQ ID NO: 258:
20
21 SEQUENCE CHARACTERISTICS:
22     LENGTH: 955 base pairs
23     TYPE: nucleic acid
24     STRANDEDNESS: single
25     TOPOLOGY: linear
26 MOLECULE TYPE: protein
27
28 FEATURE:
29     NAME/KEY: CDS
30     LOCATION: 66..944
31
32 FEATURE:
33     NAME/KEY: misc_feature
34     LOCATION: residues 1-65
35     OTHER INFORMATION: /label= EcoRI
36     /note="residues 1-65 comprise EcoRI site to beginning of pel
37     B."
38
39 FEATURE:
40     NAME/KEY: misc_feature
41     LOCATION: AA 1-22
42     OTHER INFORMATION: /label= pel B
43     /note="pel B is the leader sequence from the pectate lyase
44     gene of Erwinia carlotovora."
45
46 FEATURE:
47     NAME/KEY: misc_feature
48     LOCATION: AA 23-273
49     OTHER INFORMATION: /label= "gelonin"
50     /note="gelonin (see U.S. Patent No. US2002000638A1 5,416,202)."
```


OY	308	CCGGAAGCTACACAAATCTTGAGCGCATACGCCGTCATAGGAGCAGCAATCCCTCTGGGTA	367
Db	460	GGCGACGATATCCCTCGCTGAGAGGTGAAAGGCATATAGAGACACAACAGACTTGGGCA	519
OY	368	TAGAGAACTATTCATACCGCTCTCGGGCGCTCGTTATCGAGGGGGAGCA-----	418
Db	520	TTGAAACATTTAAGGATTGGCATCAAGAACTTGTATGAAATCGATAGACAAATTATAAC	579
OY	419	CCCGGGCCCAAGCTGTTCCCTTAATTCCTCATTCAGATGATCCGAGCGCGAGAT	478
Db	580	CAACGAGATAGCTAGTACTCTTATTGGTTGTTATCAAAATGGTGTCTGAAGCAGCTGAT	639
OY	479	TCAATCCCATCTTTTGGAGGGCGTCCCAATCATTAACAGCGGGAGTCATTTCTTCCG	538
Db	640	T-----CACTTTATTGAGAACCAAAATTAAGAAATTACTTCAACAGAGATATCGCCGG	693
OY	539	ACATGTACATCTCGAGCTGAGCTGAGACTAGTTGGGGCCCAACATCCAGCAATGTCAGAGT	598
Db	694	CGAATTAATACATCAGCCTTGAGAAATAATGGGGTAAACCTCTGTTCCAAATCCGGACAT	753
OY	599	C---TAAGCAATGGCGTTTATTAATACCATTTCGGTGGGTATATCCACCGTAACCTTCG	655
Db	754	CAGGTGCAAAATGGAATGTTTTTCGGAGGCACTGTAATGGAAACGTGCAAAATGGCAAAAAAT	813
OY	656	TGACGTTGAGCAATGTTGCGAGCTGATGCCAGCTAGGATCATGTTTGTATGTA	715
Db	814	ACTATGTACCGCAGTTGATCACTAATAAACCCAAATAGCACTCTTGAAATGTTGTCGATA	873
OY	716	GGGACGACCACTTCTTCCTC	734
Db	874	AAGATCTTAATCGGCCG	892

RESULT 10
 US-09-765-527-250
 : Sequence 250, Application US/09765527
 : Patent No. US2002000638A1
 : GENERAL INFORMATION:
 : APPLICANT: Better, Marc D.
 : TITLE OF INVENTION: Methods for Recombinant Microbial Production of
 : FUSION PROTEINS AND BPI-DERIVED PEPTIDES
 : NUMBER OF SEQUENCES: 265
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 : STREET: 6300 Sears Tower, 233 South Wacker Drive
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: United States of America
 : ZIP: 60606-6402
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/765,527
 : FILING DATE: 18-Jan-2001
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/621,803
 : FILING DATE: <Unknown>
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Borun, Michael F.
 : REGISTRATION NUMBER: 25,447
 : REFERENCE/DOCKET NUMBER: 27129/31199
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312/474-6300
 : TELEFAX: 312/474-0448
 : TELEX: 25-3856
 : INFORMATION FOR SEQ ID NO: 250:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1072 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single

Query Match	Best Local Similarity	7.8%	Score 65	DB 10	Length 1072
Matches	296	Conservative	0	Mismatches 305	Indels 18
				Gaps	3
128	GATTGTGTGTGGTGAAGTCAACCAATCAAGGGGGAGACTCGATTCACAGCGCCGCATCGACG	187			
280	GTTCCTGTTGTGTGAGCGCTTTCATTAATGACATGACAGTGTGGCGGAATPAGTATAGATG	339			
188	TTACCAATCTGTAGTGGTGGCTTACCAAGACGAGGACCAATCTTCTTTTGGCGGACG	247			
340	TTACCAAGTCTTATGTGTGTGGCTATCAAGTAAGTAAGAACAGATCTTCTTTTAAAGATG	399			
248	CACGACAGGCGGCGGAAAGGACATCTCTTACCGGCGACACAGATCTCTCTCCATTCGA	307			
400	CTCCAGATGCTGTCTTACGAAGGCTCTTCANMAACANATTAANAACAAGACTTCAATTTG	459			
308	CCGGAAGCTACACAGATCTGAGGAGATACGCCGGTCAATAGGAGCCAGATCCCTCTGGGTA	367			

Db	460	GGCGCAGCTATCCCTCGCTGGAAAGGTGAGAGGCGATATAGAGACACAGACTTGGCA	519
Oy	368	TAGAGGAACTCATTTCAATCCGCTCGGCGCTTCGTTATCCAGGCGGACGA	418
Db	520	TTGACCATTTAAGATTGGGATCATCAAGAACTTATGAAATGGGATAGCAATATATAAC	579
Oy	419	CCCGGGCCCAAGCTGTTCCTTATAATCTCATTTGATGATCTCCAGGCCGAGAT	478
Db	580	CACGGAGATAGTAGTTCCTCTATTTGGTTGTATTCAAAAGGTGTGAAAGCGCTGCAT	639
Oy	479	TCAATCCCATCTTTGGAGGGGCTCGCAATACATTTAACCGGGGAGTATTTCTTCGG	538
Db	640	T-----CACCTTATTGAGAACCAAAATTGAAATACCTTCAACAGAGATTGSCCGG	693
Oy	539	ACATGTACATCTCGAGCTGGAGCTAGTTGGGGCCACCAATCCACAGCACTCCAGAT	598
Db	694	CGAATATATACATATAGCCTTGAGATATAAAGGGGTAACTCTGTTCCAAATCCGACAT	753
Oy	599	C---TACGGATGGCGTTTAAATTAACCAATTTGGTTGGTATATCCACCGGTAACTTG	655
Db	754	CAGGTGCMAATGGAATGTTTTCCGAGGCAAGTTGAATTTGAAACGNCMAATGGCCAAAATAAT	813
Oy	656	TGACGTTGAGCAATGTTGCGGACGTGATCGCCAGCTTAGGAGATCAATGTTTGTATGTA	715
Db	814	ACTATGTCACCGCAGTTGATCAATGAATAAACCCAAAATAGCACCTCTGAAATTCGTGATA	873
Oy	716	GGGACCGACCATTTCTCTC	734
Db	874	AAGATCCCTAAATGGCGCG	892

```

RESULT 11
US-09-464-099A-2
; Sequence 2, Application US/09464099A
; Patent No. US20020168680A1
;
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOILPYRUVYLSHIKIMATE-3- PHOSPHATE SYNTHASES
; FILE REFERENCE: 1189, 0175.CNDS01. MORT:175-2
; CURRENT APPLICATION NUMBER: US/09/464,099A
; CURRENT FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Agrobacterium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (62)..(1426)
;
US-09-464-099A-2

```

[illegible]

Accession	Sequence	Position
OY	TGTACGTGTGCTTACCAACAGGGGACCAATCTCACTTTTGGGAGACCAACAG	256
OY	197	
Db	TCAATGCGGTGGATTGGATGAGAGGGGAGAAAGCTGCTGCTGTGGCGGCGCGCTTACG	1220
Db	1161	
OY	GCGCGAAGGACCTCTTTCACCGGCACCAACGATCTCTTC	298
OY	257	
Db	GCAAGGGGCTGGCAACGCTCTGGAGGGCGGCGCTGCCACCC	1262
Db	1221	

```

RESULT 12
US-09-861-696-2
; Sequence 2, Application US/09861696
; Patent No. US2002007053A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHAS
; FILE REFERENCE: 11899.0175.CUS04 MOBT-175-4
; CURRENT APPLICATION NUMBER: US/09/861,696
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/137,440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833,485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306,063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749,611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576,537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Agrobacterium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (62)..(1426)
US-09-861-696-2

```

[illegible]

```

RESULT 13
US-09-347-064-11
; Sequence 11, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5

```

CURRENT APPLICATION NUMBER: US/09/347,064A
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012.0
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 48
TYPE: DNA
ORGANISM: Viscum album
US-09-347-064-11

Query Match 4.5%; Score 38; DB 10; Length 48;
Best Local Similarity 89.1%; Pred. No. 0.0014;
Matches 41; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 730 TCCTCCGAGTGGCGCTATTGGCGGCTGTCATACGACCCGCTGTGG 775
Db 1 TCCTCGAGGTGGCTATTGGCGGCTGTCATACGACCCGCTGTATAG 46

RESULT 14
US-10-184-644-584/c
Sequence 584, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: ACIDS ENCODING THE SAME
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See file wrapper or palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 584
LENGTH: 708
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-584

Query Match 4.3%; Score 35.6; DB 9; Length 708;
Best Local Similarity 12.3%; Pred. No. 0.035;
Matches 48; Conservative 108; Mismatches 233; Indels 0; Gaps 0;

QY 125 AAGATTGTGTGGTGGAGCAATCAACAGGAGGAGAGAGTGTACAGGCGCGCATG 184
Db 663 A.RCTTYR.YG..BKC.D.N.HR.A.AC..C.S.....CCC.T.C.S.SBSBCTBMR 604
QY 185 ACCTTACATCTGTACGTGTGGCTTACCAAGCAGGACCAATCTTATTTTGGCGG 244
Db 603 RCNA.HMA.AC...B.HSDHADB..CYH.YSCA....CYSSSCHSSC..YK.NM..S 544
QY 245 ACGCAGCAGGCGGCGGAAAGGATCTTTCACGAGGAGCAGCAGAGATCTTCCAT 304
Db 543 .STSTMSDYWG..TBGA.BK..AAATTTCTB.YBA..HYTC.ST.TCC.C.YNYTSGN 484
QY 305 TCACCGGAGCTACACAGATCTGGAGCGATACCGGCTATAGGAGCAGATCCCTCTGG 364
Db 483 .STS.HS..ACS.D.BS..C.C.CYS..T..A.B..YC.SCT...C.HMSBPTH.T.S 424
QY 365 GATAGAGAACTCATTCATCCGTCGTCTGGCGCTTCGTTATTCAGAGGCGGACCGCGG 424

Db 423 T.B.G..YT.CY..T.TSD.TABB..TNBA.YT.TTCY.CS.CT.TYBRRCTH..CBS 364
QY 425 CCAAGCTCGTCCCTATATATCCATGATGATTCGAGCGCGGAGATTCATC 484
Db 363 ..Y.ATCSC.B.TCCYAMABY...NCYMTGS.STH.CHBBSR..MT....C..BH.TA 304
QY 485 CCATCTTTGGAGGCGCGCATACATY 513
Db 303 SYNTSCTNSB.CSB..A..STHTTB..Y 275

RESULT 15
US-09-864-761-13457/c
Sequence 13457, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 13457

LENGTH: 526
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC011449.3
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 16:07:41; Search time 1425.84 Seconds
(without alignments)
9507.085 Million cell updates/sec

Title: US-09-627-165E-15

Perfect score: 837
Sequence: 1 catcagacgacggcgacga.....aacccaccgtcgcacgtgta 837

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estinu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.8	5.8	405	14	T24255
2	43.6	5.2	479	12	BG789629
3	43.6	5.2	558	10	BE60625
4	42.8	5.1	811	10	BE034055
5	42.2	5.0	837	10	AM053634
6	42.2	5.0	1021	10	BE036639

7	41.6	5.0	672	10	AM053585
8	40.4	4.8	779	14	BM816277
9	40.4	4.8	913	14	BM816276
10	39.8	4.6	586	14	BQ765173
11	38.8	4.6	463	10	BE425390
12	38.8	4.6	589	9	AJ436193
13	38.8	4.6	621	10	AV836389
14	38.8	4.6	623	14	BE067985
15	38.8	4.6	626	10	BE060235
16	38.8	4.6	631	13	BE067985
17	38.8	4.6	641	14	BQ739975
18	38.8	4.6	700	9	AL505987
19	38.8	4.6	867	13	BI953600
20	38.4	4.6	420	9	AJ466017
21	38.4	4.6	540	12	BG367680
22	38.2	4.5	296	12	BG043201
23	37.4	4.5	829	17	BH526892
24	37.2	4.4	408	10	AM047216
25	37.2	4.4	561	10	BE426610
26	37.2	4.4	565	10	BE470946
27	36.4	4.3	557	14	BQ764896
28	36.4	4.3	610	14	BQ764882
29	36.4	4.3	788	12	BG310389
30	35.6	4.3	443	13	BM110994
31	35.6	4.3	545	10	BE511221
32	35.6	4.3	686	9	AI890522
33	35.6	4.3	993	10	BE035039
34	35.4	4.2	699	13	BI684857
35	35	4.2	385	13	BM136974
36	35	4.2	1094	11	AY105813
37	34.8	4.2	453	17	CNS03975
38	34.8	4.2	611	14	BQ490182
39	34.6	4.1	565	13	BI396105
40	34.6	4.1	620	9	AJ451296
41	34.6	4.1	661	9	AJ452361
42	34.6	4.1	678	9	AJ450811
43	34.4	4.1	368	14	BQ171562
44	34.2	4.1	888	12	BG369039
45	34	4.1	186	10	BA416990

ALIGNMENTS

RESULT 1
T24255
LOCUS 405 bp mRNA linear EST 28-JUL-1995
DEFINITION csl345 lambdaZAPST Rictinus communis CDNA clone pcrs1345, mRNA
ACCESSION T24255
VERSION T24255.1 GI:689074
KEYWORDS EST.
SOURCE castor bean.
ORGANISM Rictinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Rictinus.
1 (bases 1 to 405)
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)
Contact: Somerville CR
Carnegie Institution
Carnegie Institution, 290 Panama St, Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.

FEATURES
source Location/Qualifiers
1..405
/organism="Rictinus communis"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrs1345"

Query Match 5.18; Score 42.8; DB 10; Length 811;

110 CCGTTCGGATCGCAAGATTTCTGTTCGTGAACACTCACCATTACGGGGGAGACACTCGA 169

[illegible]

RESULT 7	AM053585	672 bp	mRNA	linear	EST 20-FEB-2001
LOCUS	U30-214375	Ice plant Lambda Uni-Zap XR expression library, 30 hours			
DEFINITION	NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-2143				
ACCESSION	AM053585				
VERSION	AM053585.1	GI:5916778			
KEYWORDS	5', mRNA sequence.				
SOURCE	EST.				
ORGANISM	common iceplant.				
	Mesembryanthemum crystallinum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;				
	Caryophyllales; Caryophyllales; Aizoaceae; Mesembryanthemum.				
REFERENCE	1 (bases 1 to 672)				
AUTHORS	Cushman, J.C.				
TITLE	An expressed sequence tag database for the common ice plant,				
	Mesembryanthemum crystallinum				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Cushman JC				
	Department of Biochemistry				
	University of Nevada				
	MS200, Reno, NV 89557-0014, USA				
	Tel: 775-784-1918				
	Fax: 775-784-1650				
	Email: jcushman@unr.edu				
	PCR Primers				
	FORWARD: T7				
	BACKWARD: T3				
	Plate: L30-22	row: D	column: 11		
	Seq primer: T3				
	High quality sequence stop: 450.				
FEATURES	Location/qualifiers				
source	1..672				
	/organism="Mesembryanthemum crystallinum"				
	/db_xref="taxon:3544"				
	/clone="L30-2143"				
	/clone.lib="Ice plant Lambda Uni-Zap XR expression library				
	/30 hours NaCl treatment"				
	/tissue.type="leaf, 30 h 0.4M NaCl"				
	/dev.stage="Six week old"				
	/note="vector: lambda Uni-Zap XR, Bluescript SK-; site_1:				
	EcoRI; site_2: XhoI"				
BASE COUNT	209 a 182 c 133 g 148 t				
ORIGIN					
Query Match	5.0%; Score 41.6; DB 10; Length 672;				
Best Local Similarity	54.0%; Pred. No. 0.58;				
Matches 108; Conservative	0; Mismatches 89; Indels 3; Gaps 1;				
QY	15 CGAGCAATATTTTCGGTTCATCAACGCTTCCGAGATTATGTCACAGCGGAGACTTTTC 74				
DB	228 CGACACATATTCACAGTTCATTCGCGATCTCCAACTGAATATTCTCCGCGACACACCG 287				
QY	75 CAATGAGATTAACACTCTGGGTGAGTCTAGATCCCGCTGCGATGCGCAAAATTTGT 134				
DB	288 ATGTCAATATCC---CGTGAACAGGGGGAACCGCAACCATCATCAATGACCAACGATTTGT 344				
QY	135 GTTGGTGAACTCAACCAATCAGGGGGAGACGTGATCAGCGCGGCATGCACTTACCA 194				
DB	345 CCTGTCACCTCCGCAAAACCCCTGGCAAAAACCATCACCCTCGCATGCACTGACAGA 404				
QY	195 TCTGACGTGGGGCTTACC 214				
DB	405 CGCTACGTCGTGCGCTATC 424				
RESULT 8					
LOCUS	BM816277	779 bp	mRNA	linear	EST 05-MAR-2002
DEFINITION	HB106E08_SK.ab1 HB Hordeum vulgare cDNA clone HB106E08_SK.ab1				

SECRET

```
/tissue_type="root"
```

```

/1ab_host="DH10B"
/note="Vector: pSPORT1; Site.1: Sal I; Site.2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old waterlogged barley
plants. Developed as part of the barley transcriptome
resources of BSRRC/SEERAD funded cereal IGF (Investigating
Gene Function) project."
BASE COUNT      81 a      248 c      170 g      87 t
ORIGIN
Query Match      4.8%; Score 39.8; DB 14; Length 586;
Best Local Similarity 52.8%; Pred. No. 1.8;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 164 ACTGATACAGCGCCGATGACGTTACCAATGTACGTGGGTACCAAGCAGCGG 223
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 376 ACATGCCCCCGCCCTTGTGCGCAAGACCCCGGTGGCGCTCTGACGCGG 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 ACCAATCTACTTTTTCGCGACGACCAAGCGCGGAAAGCATCTTTCACGGCA 283
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CCTCCTTCCGGGCTCTTGTGACACCTCGCTCGCGACAGACGACCTTTCACGGCA 495
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 CCACCAAGATCCTCTCTCCATTCACGGAAGCTACACATCT 326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 CCTTCATGCCCTCCACCGACCTACCGCGGCTACCGCGTCT 538
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BE425390      463 bp      mRNA      linear      EST 24-JUL-2000
LOCUS      WHE313_H06.H06ZS wheat unstressed seedling shoot cDNA library
DEFINITION      Triticum aestivum cDNA clone WHE313_H06_H06, mRNA sequence.
ACCESSION      BE425390
VERSION      BE425390.1 GI:9423233
KEYWORDS      EST.
SOURCE      bread wheat.
ORGANISM      Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 463)
AUTHORS      Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Iazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
COMMENT      The structure and function of the expressed portion of the wheat
genomes
JOURNAL      Unpublished (2000)
CONTACT      Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@wpr.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stragene SK primer.
FEATURES
Source      location/Qualifiers
1..463
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE313_H06_H06"
/clone_lib="wheat unstressed seedling shoot cDNA library"
/tissue_type="Etiolated shoot"
/dev_stage="Five day old seedling"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site.1: EcoRI; Site.2: XhoI; Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Shoots were
harvested. The tissue, total RNA, and poly(A) RNA were

```

```

prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give Bluescript phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
BASE COUNT      68 a      207 c      103 g      85 t
ORIGIN
Query Match      4.6%; Score 38.8; DB 10; Length 463;
Best Local Similarity 54.1%; Pred. No. 3;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 173 CGCGCCGATGACGCTTACCAATCTGTACGTGGCTTACCAAGCAGCGCAATCT 232
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 CGGGCCCTTATGCGGGCGGACCCGCGGTGTGCGCGCTCGACGCCAAGAGCTTGC 66
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 ACTTTTGGCGAGCAGCAGCAGCGCGGGAAGGACATCTCTACCGCGGCCACCGAT 292
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 CGGTCTCTTCGACGCTGACAGAGTGGAGAAAGAACCTTTCACCGGACCTATATGC 126
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 CCTCTCCCATTCACCGGAAGCTAC 318
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 CCTCACCTCCTTCACCGGCGGCTTC 152
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
AJ436193      589 bp      mRNA      linear      EST 15-MAR-2002
LOCUS      AJ436193 S00007 Hordeum vulgare cDNA clone S0000700017B08F1, mRNA
DEFINITION      sequence.
ACCESSION      AJ436193
VERSION      AJ436193.1 GI:19524645
KEYWORDS      EST.
SOURCE      Hordeum vulgare.
ORGANISM      Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE      1 (bases 1 to 589)
AUTHORS      Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
COMMENT      Unpublished (2002)
CONTACT      Schulman AH
Institute of Biotechnology
University of Helsinki
P.O.Box 56 (Viikinkaari 6A), University of Helsinki FIN-00014,
Finland.
FEATURES
Source      location/Qualifiers
1..589
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone="S0000700017B08F1"
/clone_lib="S00007"
/dev_stage="Shoot"
/note="2,3,-4-days after germination"
BASE COUNT      102 a      237 c      147 g      103 t
ORIGIN
Query Match      4.6%; Score 38.8; DB 9; Length 589;
Best Local Similarity 54.1%; Pred. No. 3.5;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 173 CGCGCCGATGACGCTTACCAATCTGTACGTGGCTTACCAAGCAGCGCAATCT 232
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CGGGCCCTTATGCGGGCGGACCCCGGTGTGCGCGCTCGACGCCAAGAGCTTGC 351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 233 ACTTTTGGCGAGCAGCAGCAGCGCGGGAAGGACATCTCTTCACCGGCGCACCAAGAT 292
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 CGGTCTCTTCGACGCTGACAGAGTGGAGAAAGAACCTTTCACCGGCGCACCTATATGC 411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 293 CCTCTCCCATTCACCGGAAGCTAC 318
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 412 CCTCCACCTCCTCACCGGCGGCTTC 437

RESULT 13
AV836389 621 bp mRNA EST 22-JUN-2001
LOCUS AV836389 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
DEFINITION vulgare seedling leaves second leaf stage Hordeum vulgare subsp.
vulgare CDNA clone basdlic19, mRNA sequence.

ACCESSION
VERSION AV836389
KEYWORDS GI:14528478
SOURCE EST.
ORGANISM Hordeum vulgare subsp. vulgare.
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.

REFERENCE
AUTHORS Sato, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2001)
COMMENT Contact: Kazuhiko Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazasato@rib.okayama-u.ac.jp,
URL: http://www.rib.okayama-u.ac.jp/barley/
Sato, K., Saitsho, D., Takeda, K., Shint, T. and Kohara, Y. Direct
submission:
database: http://www.shigen.nig.ac.jp/barley/Barley.html.
Location/Qualifiers
1..621
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="basdlic19"
/clone_lbp="K. Sato unpublished cDNA library: Hordeum
vulgare subsp. vulgare seedling leaves second leaf stage"
/russue_type="seedling leaves"
/dev_stage="second leaf stage" 110 t

BASE COUNT 109 a 251 c 151 g 110 t

ORIGIN

Query Match 4.6% Score 38.8; DB 10; Length 621;
Best Local Similarity 54.1%; Pred. No. 3.6; Mismatches 67; Indels 0; Gaps 0;
Matches 79; Conservative 0;

QY 173 CGCGGCCATCGACATTGATGTAGTGCATGGTGTTACTCAAGCAGGCGACCAATCCT 232
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 280 CGGGCCCTCTTAATGCGGGCGAGCCCCCGCTGCGTGCCTGACGCCAAGAAGCTTCC 339
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 233 ACTTTTGGCGAGCACCAGACGCGCGGGAAGACATCTTTCACCGGCGACACCATGAT 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 340 CCGTCTCTTCACGCTCACCAAGTGCAGAAAGAACCTTTCACCGGCGACTTAGTC 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 293 CCTCTCCCATTCACCGGAAGCTAC 318
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 400 CCTCCACCTCCTCACCGGCGGCTTC 425

RESULT 14
BQ67985 623 bp mRNA linear EST 26-JUL-2002
LOCUS BQ67985 EBRc08_S0010_J11_R root, 3 week, drought-stressed, cv Optic, EBRc08
DEFINITION Hordeum vulgare CDNA clone EBRc08_S0010_J11_5', mRNA sequence.
ACCESSION BQ67985
VERSION BQ67985.1 GI:21976459
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.

REFERENCE	1 (bases 1 to 623)
HEADLEY, P., LIU, H., CALDWELL, D., MCCALLUM, N., MUDIE, S., CARDLE, L., RAMESY, L., MACHRAY, G., MARSHALL, D.F.M. and WAUGH, R.	
TITLE	Development of Barley Transcriptome Resources
JOURNAL	Unpublished (2001)
COMMENT	Contact: Waugh R, Marshall DF Genome Dynamics/Computational Biology Scottish Crop Research Institute Invergowrie, Dundee, DD2 5DA, Scotland, UK Tel: 00 44 1382 562731 Fax: 00 44 1382 562426 Email: estescr1.sarl.ac.uk. Location/Qualifiers
FEATURES	
source	1..623 /organism="Hordeum vulgare" /cultivar="Optic" /db_xref="taxon:4513" /clone="EBrc08_S0010_111" /clone_id="root", 3 week, drought-stressed, cv Optic, EBrc08" /tissue_type="root" /dev_stage="3 week" /lab_host="DH10B" /note="Vector: pSPOR1; Site1: Sal I; Site2: Not I; Non-normalised library, directionally cloned into pSPOR1. Derived from roots of 3 week old drought stressed barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."
BASE COUNT	109 a 249 c 156 g 109 t
ORIGIN	
Query Match	4.6%; Score 38.8; DB 14; Length 623;
Best Local Similarity	54.1%; Pred. No.3.6; 67; Indels 0; Gaps 0;
Matches	79; Conservative 0; Mismatches 67;
QY	173 CGCGCCGACATCAGCTTACCATCTGTACGTTGGTGTACCAAGCAGCGACCATCT 232 DB 285 CGGGCCCCCTTACATGGCGCGCGACCCCGCGTGTGCTGCAGCCCAAGACTTCC 344
QY	233 ACTTTTTCGCGCAGCAGCAGCAGCGCGCGGAAAGGACATCTTTCACCGCGCACCAAT 292
DB	345 CGGTGCTCTTTCAGACGTCACCAAGGTGAGAGAAACCTTTCACCGGACATCAATGC 404
QY	293 CCTCTCTCCCATTCACCTCGGAGCTAC 318
DB	405 CCTCAGCTCTCCTCACCGCGGCTTC 430
RESULT 15	
BE060255	626 bp mRNA linear EST 22-OCT-2001
LOCUS	HVSMERG001M01f Hordeum vulgare pre-anthesis spike EST library
DEFINITION	HVCNNA0008 (white to yellow anther) Hordeum vulgare cDNA clone HVSMERG001M01f, mRNA sequence.
ACCESSION	BE060255
VERSION	BE060255.2 GI:13154152
KEYWORDS	EST.
SOURCE	Hordeum vulgare.
ORGANISM	Hordeum vulgare. Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae; Triticeae; Hordeum. 1 (bases 1 to 626) Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Fritsch, D., Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Close, S.J., Oates, R. and Main, D. Development of a genetically and physically anchored EST resource for barley genomics: Morex pre-anthesis spike cDNA library Unpublished (2001) On Jun 9, 2000 this sequence version replaced gi:8404905. Contact: Wing RA Clemson University Genomics Institute
TITLE	
JOURNAL	
COMMENT	

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu

Total hg bases = 236
Seq primer: AATTACCTCCTAAGG
High quality sequence stop: 435.
Location/Qualifiers

FEATURES

source

1. .626

/organism="Hordeum vulgare"

/cultivar="Morex"

/db_xref="taxon:4513"

/clone_id="HVSME9011M01f"

/clone_lib="Hordeum vulgare pre-anthesis spike EST library

HVCDA0008 (white to yellow anther)"

/tissue_type="pre-anthesis spike"

/lab_host="SOLR"

/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;

Plants were grown in the greenhouse at the University of

California, Riverside (Fenton, SJ Close, TJ Close). Whole

spike with awns trimmed were collected at white, green and

yellow anther stages (Fenton). Total RNA was prepared from

each pool, equal quantities of all three RNA pools were

combined, poly(A) RNA was purified from the mixture, one

primary unamplified cDNA library was made, and 1 million

plu were in vivo excised to give Bluescript SK(-) cDNA

phagemids. These steps were performed in the TJ Close lab

(Choi) at the University of California, Riverside.

Phagemids were plated and picked at the Clemson University

Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins

and Wing) plasmid DNA preparations, DNA sequencing and

sequence analysis were performed at CUGI (Wing, Yu, Frisch

, Henry, Simmons, Oates, Rambo, Main). The sequence has

been trimmed to remove vector sequence and contains a

minimum of 100 bases of phred value 20 or above. For more

details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order

this clone see <http://www.genome.clemson.edu/orders> Also

see Close TJ, Wing R, Kleinholz A, Wise R (2001)

Genetically and physically anchored EST resources for

barley genomics. Barley Genetics Newsletter 31:29-30.

(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 99 a 261 c 141 g 116 t 9 others

ORIGIN

Query Match

Best Local Similarity 4.6%; Score 38.8; DB 10; Length 626;

Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

OY 173 CGGCGCCATCGAGCTTACCAATCTGTACGTGTGCGCTTACCAAGCAGGACCAATCCT 232

DB 113 CGGCGCCCTTTCATGCGCGCGCCGACCCCGCGGTGCGCGCTCGACGCCAAGAGCTTCC 172

OY 233 ACTTTTTCGCGAGCAGCAGCAGCGCGGGAAGGATCTTTCACCGGACACCAAGAT 292

DB 173 CCGTGTCTTTCGAGTCACCAAGGTGAGAGAAGAACCTTTCACCGGACACTTACATGC 232

OY 293 CCTCTGTCCCAFTACCGGGAAGCTAC 318

DB 233 CTTCCACTTCCCTCACCGGCGGCTTC 258

Search completed: March 26, 2003, 19:41:29
Job time : 1434.84 secs

Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schmale@iastate.edu

Individual basecall and confidence value were assigned using the phred software,
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm> >#br/rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (<http://www.tigr.org/softlab/>).

Lucy parameters were set to ensure an overall trimmed quality of 97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.

PCR Primers

FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)

BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

FEATURES

source

1. .550
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST606-C10"
/clone_lib="ISUM5-RN"
/tissue_type="mixed"
/lab_host="DH10B"

/note="vector: pT73PAC; Site_1: EcoRI; Site_2: NotI; Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG), Mixed mature tissues (17, 21, 38, 69, 77 DAG), kernels (3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65 DAG), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0 cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated first ear, ear shank, etiolated seedlings, callus, Cycloheximide-treated callus, Anaerobic treated seedlings, NAA (a-Naphthalene acetic acid)-treated seedlings, Kinetin-treated seedlings, ACP (1-aminocyclopropane-1-carboxylic acid)-treated seedlings, Brassinolide-treated seedlings, ABA (Abscissic acid)-treated seedlings, GA (Gibberellic acid)-treated seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5', AACTGAGAGATTCGGCCCGCAGCAATTTTCTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT73PAC vector. The library then went through one round of normalization to Cot value of 5 based on the methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."

BASE COUNT 140 a 150 c 125 g 135 t
ORIGIN

Query Match

Best Local Similarity 52.1%; Score 33.8; DB 14; Length 550;

Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY	150	CCAAATFAAGATTGCAGATTCCGGCGGTAACTTGTGACGNTGAGCAATGTCGGAGCT	209
DB	521	CCGGATCAGACAGATTAATGCGACGACAAAGTTGGACCTGGGCAATTCATCAGAAAT	462
QY	210	GATCTCCAGCTTGGCAGTGTGTTGCAATGCAATGCGGCAATTCCTCTCGA	269
DB	461	CGTCTCCAGACATGGGACCAAGATGATCGAGGCTGTCGAGACAGCCGTCGGCCACGA	402
QY	270	CCACCTTGGCGCTCTCCTA	291
DB	401	CGACATCAACTCGTCTCTGCTA	380

Search completed: March 26, 2003, 19:41:20
Job time : 618.156 secs

QY 193 AGCAATGTTGCGACGTGATC 213
 Db 311 GACAAATGACCTCATGCTGATC 331

RESULT 13
 CNS01XEV/c
 LOCUS
 DEFINITION CNS01XEV 855 bp DNA linear GSS 12-MAY-2000
 Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone
 202622 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION AL171616
 VERSION AL171616.1 GI:7809673
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 855)
 AUTHORS Roest-Crollius,H., Jallou,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Queller,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

TITLE Unpublished
 JOURNAL 2 (bases 1 to 855)
 AUTHORS Roest-Crollius,H., Jallou,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billaud,A., Queller,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 AUTHORS 3 (bases 1 to 855)
 TITLE Genoscope.
 JOURNAL Direct Submission
 COMMENT Submitted (12-Apr-2000)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES
 source
 1..855
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="202622"
 /clone_1lb="G"
 /note="Genoscope sequence ID : C0AG202BD11SP1-end :
 PUC-Or1"

BASE COUNT 136 a 146 c 286 g 243 t 44 others

ORIGIN

Query Match 9.5%; Score 34; DB 17; Length 855;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 67; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 232 TTGTTGCAATGCAGTGGTGGCCATTCTCTCGACACACCTTGGCCGCTGCTCCTA 291
 Db 291 TAGTACGTACTACACACATCTCATGATCCCTCATGTACACCTTGCAGCGCATCACTA 232

QY 292 AGGTCCGTCGTGATGCGGCCACACATGTCACTCACTTNTTTCGAACCCACGCTGGCG 351
 Db 231 CTGCCCTTATGATGACCTCATGATGTGCGTCAGACGATTCGATGCCAAGGAGCGC 172

QY 352 ATC 354
 Db 171 CTC 169

RESULT 14
 BF526014/c

LOCUS BF526014 1078 bp mRNA EST 11-DEC-2000
 DEFINITION 602070191P1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4212886
 5', mRNA sequence.

ACCESSION BF526014
 VERSION BF526014.1 GI:11613375
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1078)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LIA9784 row: f column: 23
 High quality sequence stop: 660.

FEATURES
 source
 1..1078
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4212886"
 /clone_1lb="NCI_CGAP_Brn64"
 /tissue_type="gillblastsoma with EGFR amplification"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Brain; Vector: pCMV-Sport6; Site:1; NotI;
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.57 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 282 a 236 c 350 g 149 t 1 others

ORIGIN

Query Match 9.5%; Score 34; DB 12; Length 1078;
 Best Local Similarity 56.1%; Pred. No. 14;
 Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 191 TGACCAATGTTGCGGACGTGATCTCCAGCTTGGGATCATGTTTGAATGAGTGGTC 250
 Db 810 TGAACACAGTTTCGCGGCTCTTGATCCGGCTGTGATGCTCCGATTCCTCGTGGC 751

QY 251 GGCATCTCTCTCTCTGACACACCTCGCGCTGCTCCCAAGTCCGTCGTGG 304
 Db 750 GCTCTCTCGCATCTGATGATCTGTGCTGCTTACCATGCACCATGGGG 697

RESULT 15
 BO538836/c
 LOCUS
 DEFINITION BO538836 550 bp mRNA linear EST 12-JUN-2002
 MEST606-C10.T3 ISUM5-RN Zea mays cDNA clone MEST606-C10 3', mRNA
 sequence.

ACCESSION BO538836
 VERSION BO538836
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade: Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 550)
 AUTHORS Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
 TITLE Expressed Sequence Tags from B73 Maize: various stages and tissues
 JOURNAL including seedlings treated with a variety of hormones
 COMMENT Unpublished (2001)
 Contact: Patrick S. Schnable
 Schnable Laboratory

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 879)	Mammalia: Eutheria: Primates, Catarrhini, Homnidae; Homo.	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC). Unpublished (1999)	Contact: Robert Strausberg, Ph.D.

BASE COUNT ORIGIN	284 a	209 c	182 g	204 t
----------------------	-------	-------	-------	-------

RESULT 9	485 bp	mRNA	linear	EST 30-JUN-2000
LOCUS	AJ284561			
DEFINITION	4A3B-P2A9-F Anopheles gambiae immune competent 4A3B Anopheles			
ACCESSION	AJ284561			
VERSION	AJ284561.1	GI:6932440		
KEYWORDS	EST.			
SOURCE	African malaria mosquito.			
ORGANISM	Anopheles gambiae			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;			
REFERENCE	1 (bases 1 to 485)			
AUTHORS	Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,			
	Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.			
	and Kafatos, F.C.			
TITLE	Anopheles gambiae pilot gene discovery project: identification of			
	mosquito innate immunity genes from expressed sequence tags			
	generated from immune-competent cell lines			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (12),	6619-6624	(2000)
MEDLINE	203009950			
COMMENT	Contact: Dimopoulos G			

FEATURES
source

BASE COUNT	95 a	111 c	159 g	120 t
ORIGIN				

```

RESULT 10
CNS07HYF/C
LOCUS      586 bp      DNA      linear      GSS 03-OCT-2001
DEFINITION Anopheles gambiae GSS T7 end of clone 31J01 of library Notredame1
            from strain PEST of Anopheles gambiae (African malaria mosquito),
            genomic survey sequence.
AL611673
ACCESSION AL611673.1  GI:15963096
VERSION   GSS.
KEYWORDS  African malaria mosquito.
SOURCE    Anopheles gambiae
ORGANISM  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
            Anopheles.
REFERENCE 1 (bases 1 to 586)
AUTHORS   Genoscope.
TITLE     Direct Submission
JOURNAL   Submitted (01-OCT-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 Evry cedex - FRANCE (E-mail : seq@genoscope.cns.fr)
COMMENT   Web : www.genoscope.cns.fr)
REFERENCE 2 (bases 1 to 586)
AUTHORS   Roth,C.W., Brey,P.T., Ke,Z. and Collins,F.H.
TITLE     Direct Submission
JOURNAL   Submitted (01-OCT-2001) BBMI, Institut Pasteur, 25, rue du Dr.
            Roux, Paris 75015, France
COMMENT   This clone is from an A. gambiae BAC library provided by F.H.
            Collins and sequenced by Genoscope in collaboration with the
            Laboratory of Biochem. and Biol. Molec. of Insects, Institut
            Pasteur.

FEATURES
     source              Location/Qualifiers
         1..586
            /organism="Anopheles gambiae"
            /strain="PEST"
            /db_xref="taxon:7165"
            /clone="31J01"
            /clone_lib="Notredame1"
            /note="end : T7"

```


Db 489 CAGCGGGGTCCTTAAGCTTGAAGATTCTGGGCTCGAGAGATCAGGTGTTTCT 430
 QY 263 CTTCTGACCACTTCCTGGCGCTGCTCTTAAGT 295
 Db 429 TTGTGTTGCTGCTTCCTCCCTCTCTCTCTGAGCT 397

RESULT 4

BI474176 554 bp mRNA linear EST 27-AUG-2001

LOCUS f35h12.y3 zebrafish gridded kidney Danio rerio cDNA clone 4745350

DEFINITION 5', similar to TR:09VPF8 09VPF8 CELL-ADHESION PROTEIN ALPHA

ACCESSION BI474176
 VERSION BI474176
 KEYWORDS EST

SOURCE zebrafish.
 ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 554)

REFERENCE

AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,R., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepcoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE Washu zebrafish EST Project 1998
 JOURNAL Unpublished (1998)

COMMENT

Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrflsh@wustl.edu

CDNA library Preparation: Leonard Ira Zon DNA Sequencing by:
 Washington University Genome Sequencing Center Clone distribution:
 Genome Systems, St. Louis, Missouri (web address:
 www.genomesystems.com) (email contact: info@genomesystems.com) and
 Research Genetics, Huntsville, Alabama (web address: www.resgen.com
) (email contact: info@resgen.com) and
 ResourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
 www.rzpd.de)

FEATURES
 source High quality sequence stop: 343.
 Location/Qualifiers
 1..554
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="4745350"
 /clone_lib="zebrafish gridded kidney"
 /sex="mixed"
 /tissue_type="kidney pooled from 300 wild type adults"
 /lab_host="XIOIR"
 /note="Organ: Kidney; Vector: pBR-CMV; Site:1: EcoRI; Site:2: XhoI; Oligo dT cDNA library constructed from mRNA pooled from pooled kidney tissue from 300 adult zebrafish."

BASE COUNT 168 a 122 c 142 g 122 t
 ORIGIN

Query Match 10.1%; Score 36; DB 13; Length 554;
 Best Local Similarity 53.2%; Pred. No. 2.3;
 Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 214 TCCAGCTTGGCGATCATGTTGTTGATGAGTGGCGCATTTCTCTTCGACAC 273
 Db 415 TCCAGTGGGCTGCCAATTCATGAGCCCTCTCTTTGTTAGATGGA 356
 QY 274 CCTTCGCGGCTGCTCTTAAGTCCCTGCGATGCGGCAAGATGTCACCTGCAC 333
 Db 355 CCTTTGTTGCTGAGATTACAGTGTGGTACCTGGGTTACCAATGGCTCAGCAG 296

QY 334 TCCGAACCCAGCGTCGCATC 354
 Db 295 TCCGACGACAGCGTACGATC 275

RESULT 5

BI885867 749 bp mRNA linear EST 12-OCT-2001

LOCUS ZF637-1-000122 zebrafish shield stage whole embryo cDNA library

DEFINITION MPMGP637 Danio rerio cDNA clone MPMGP637_6F9;MPMGP637F096 5', mRNA sequence.

ACCESSION BI885867
 VERSION BI885867
 KEYWORDS EST

SOURCE zebrafish.
 ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 749)

REFERENCE

AUTHORS Clark,M., Aanstad,P., Hennig,S., Johnson,S.L. and Lehrach,H.
 EST sequencing of a zebrafish shield stage cDNA library normalised by oligonucleotide fingerprinting
 Unpublished (2001)

TITLE Contact: Hennig S
 JOURNAL Laboratory 123, Dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380

COMMENT

Email: hennig@molgen.mpg.de
 5' EST sequencing of clones from a zebrafish shield stage library, normalised from 55,000 starting clones by oligonucleotide fingerprinting
 High quality sequence stop: 749.
 Location/Qualifiers
 1..749
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="MPMGP637-6F9;MPMGP637F096"
 /clone_lib="zebrafish shield stage whole embryo cDNA library MPMGP637"
 /tissue_type="whole embryo"
 /dev_stage="shield stage, 6 hrs post-fertilisation"
 /lab_host="E.coli, XL1 blue MRP"
 /note="Vector: pSport1; Site:1: NotI; Site:2: SalI; oligo-dT-NotI primed, SalI adaptor, directionally cloned, library normalised by oligonucleotide fingerprinting"

BASE COUNT 182 a 236 c 150 g 181 t
 ORIGIN

Query Match 10.1%; Score 36; DB 13; Length 749;
 Best Local Similarity 53.2%; Pred. No. 2.7;
 Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 214 TCCAGCTTGGCGATCATGTTGTTGATGAGTGGCGCATTTCTCTTCGACAC 273
 Db 409 TCCAGTGGGCTGCCAATTCATGAGCCCTCTCTTTGTTAGAGGAC 350
 QY 274 CCTTCGCGGCTGCTCTTAAGTCCCTGCGATGCGGCAAGATGTCACCTGCAC 333
 Db 349 CTTTGTGTTGATGAGTTTACAGATGTGTGACCTGGGTTACCAATGGCTCAGCAG 290
 QY 334 TCCGAACCCAGCGTCGCATC 354
 Db 289 TCCTCACGTAGCGTACGATC 269

RESULT 6
 LOCUS BG880257
 DEFINITION BG880257 491 bp mRNA linear EST 13-MAR-2002
 ib68B04.Y1 Melton Amplified Mouse E10 5 12 5 Pancreas 1 M10S1-A Mus
 musculus cDNA clone IMAGE:5651670 5' similar to TR:09R077 09R077

BASE COUNT	ORIGIN
207 a	333 c 360 g 194 t

Query Match	12.4%	Score 44.2;	DB 11;	Length 1094;
Best Local Similarity	52.8%;	Pred. No. 0.0085;		
Matches 94;	Conservative 0;	Mismatches 84;	Indels 0;	Gaps 0;

```

Qy      68  ACATGTACATGCTCGAGCTGGAGACAGTGGGGTTCGACATCCACCAGTCCAGCAGT 1
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      809  ACCCGAGCCTGCGGGGGCTCCAGGCGGGCTGGCCGTGGCTCTCCGGCGCCGTCCAGCAGT 8

```

QY 128 CCAGGATGGCATTTTAAATACCAATAAGATTGCAGATTCCGCCGTACTTTGTGA 1
||| ||| | ||| ||| ||| ||| |||
Db 869 CCTACCAAGGGCGCCTTCGCCGCCCGGTCCTGTCTCCAGCGAACAACGGCAAGTGATGC 9

QY 188 CGNTGACCATGTTCGGCAGCTGATCTCCAGCTTGGCGAATCATGTGTGGAATGCAG 245
+
Db 929 GGGTGACAACAGTGGCGTAGCCGGTCCCCCCTCGTGTCTTCCITGGAGCACGGAAGCTG 986

| RESULT 2 | BP642076/c | 503 bp | mrna | linear | EST 19-DEC-2000 |
|------------|-------------------------------|--------|-----------|----------|-----------------------|
| LOCUS | BP642076 | | | | |
| DEFINITION | BF642076 | Insect | hemolymph | Medicago | truncatula cDNA clone |
| | NE031F051NF1046 | | | | |
| | NE031F051N 5', mRNA sequence. | | | | |

| | |
|-----------|--------------------------------------|
| ACCESSION | BF6420/6 |
| VERSION | BF642076.1 |
| KEYWORDS | GI:11906234 |
| SOURCE | EST. |
| ORGANISM | barrel medic.
Medicago truncatula |

| | |
|-----------|--|
| REFERENCE | 1 (bases 1 to 503) |
| AUTHORS | Korh  , K., Scott, A.D., Harris, A.R., Gonz  les, R.A., Bell, C.J., Florencia, H.R., Inman, J.T., Weller, J.W. and May, G.D. |
| TITLE | Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula insect herbivory library |
| JOURNAL | Unpublished (2000) |
| COMMENT | Contact: Korh  , K. |

Dept. of Plant Pathology
University of Arkansas
217 Plant Science Building, Fayetteville, AR 72701, USA
Tel: 501 575 5191
Fax: 501 575 7601
Email: khorthe@comp.uark.edu
Insert Length: 503 Std Error: 0.00
Plate: 031 row: F column: 05
Seq primer: TCACACAGCAACACCTATGAC.
Location/Qualifiers

```
source
1..503
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF031F051N"
/clone_1lb="Insect herbivory"
/tissue_type="local and systemic leaves"
```

BASE COUNT 184 a 100 c 96 g 117 t 6 others

| | | | | |
|-----------------------|--------|-----------------|--------|----------------|
| Query Match | 11.0%; | Score 39.2; | DB 12; | Length 503; |
| Best Local Similarity | 53.0%; | Pred. No. 0.22; | | |
| Matches | 80; | Conservative | 0; | Mismatches 71; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

[illegible]

245 GTGGTCGCGCAATCTTCCCTCTCTGTGACCACCCCTTGCCGCCTGCTCCTAGGTCGCCGTGG 304
| | | | | | | | | | | | | | | |
205 TTCTTTCCAACCTTCACCTTTCACAATCCCTTAACCCCTTCTTCCCTCTTCTTCCATCATCC 146

305 ATGGGGCAACGATGTACCTGCACCTNTTTC 335
 11 1 1 1 1 1 1 1 1
 145 ATCTTCTAGCTCCTTCACCTTTCATTCCNMC 115

RESULT 3
GG608460/c

| | | | | | |
|------------|---|-------------|------|--------|-----------------|
| LOCUS | BG608460 | 599 bp | mRNA | linear | EST 1/-APR-2000 |
| DEFINITION | 307054 MARC 1P1G Sus scrofa CDNA 5', mRNA sequence. | | | | |
| ACCESSION | BG608460 | | | | |
| VERSION | BG608460.1 | GI:13658438 | | | |

| KEYWORDS | EST. |
|----------|--|
| SOURCE | pig.. |
| ORGANISM | Sus scrofa |
| | Fukarovita: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; |

REFERENCE

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 599)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Grosse,W.M., Bennett,G.A., Lacroix,W.W.

TITLE design and use of two pooled tissue normalized cdna libraries for EST discovery in swine

JOURNAL COMMENT
Unpublished (2000)
Contact: Smith TPL
USA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD. AGGAACACCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG
plate: 94 row: M column: 21
Seq primer: ATTTAGGTGCACACTATAG

```

FEATURES
source
location/Voucher#s
1. 599
/organism="Sus scrofa"
/db_xref="taxon:9823"

```

```

/clone_lib="MARC IP1G"
/tissue_type="pooled"
/lab_host="DH10B"
/notes="Vector: pCMV SPOR6; Site_1: NotI; Site_2: SalI;

```

| BASE COUNT | LIBRIGIN | LIBRARY | EMBRIGOS | POOLED |
|------------|----------|---------|----------|--------|
| 152 | a | 184 | c | 134 |
| | | | | g |
| | | | | 129 |
| | | | | t |

| | | | | |
|-----------------------|-------|----------------|-------|---------------|
| Query Match | 10.8% | Score 38.4 | DB 12 | Length 599 |
| Best Local Similarity | 52.9% | Pred. No. 0.43 | | |
| Matches | 81 | Conservative | 0 | Mismatches 72 |
| | | | | Indels 0 |
| | | | | Gaps 0 |

[illegible]

D5 549 TTGAGAAACCCAGAAATTTAAATAATCTAACCGGAACATCTGTGGTTCATCATCATTCTA
D6 203 GCGACGTGATCTCGAGCTTGCGGATCATGTTGTTGGAATGCAGTGGTGGCCATTTCCT

GenCore version 5.1.4-D5_A578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 16:07:41 ; Search time 608.156 Seconds

(without alignments)
9507.085 Million cell updates/sec

Title: US-09-627-165E-13

Perfect score: 357

Sequence: 1 gccacatccatccatccatc.....aacccaccgtgcgcacgcga 357

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 44.2 | 12.4 | 1094 | 11 | AY105813 Zea mays |
| 2 | 39.2 | 11.0 | 503 | 12 | BF642076 NFO31F051 |
| 3 | 38.4 | 10.8 | 599 | 12 | BG608460 307054 MA |
| 4 | 36 | 10.1 | 554 | 13 | B1474176 fp35h12.Y |
| 5 | 36 | 10.1 | 749 | 13 | B1885867 ZF037-1-0 |
| 6 | 35 | 9.8 | 491 | 12 | BG880257 1b68b04.Y |

| | | | | | | | |
|---|----|------|-----|------|----|-----------|---------------------|
| C | 7 | 34.8 | 9.7 | 679 | 14 | BM792140 | BM792140 K-EST0072 |
| C | 8 | 34.8 | 9.7 | 879 | 12 | BF576920 | BF576920 602134355 |
| C | 9 | 34.4 | 9.6 | 485 | 9 | AJ284561 | AJ284561 4A3B-P2A9 |
| C | 10 | 34.4 | 9.6 | 586 | 17 | CNS07HYF | AL611673 Anopheles |
| C | 11 | 34.4 | 9.6 | 632 | 13 | BM624233 | BM624233 170006874 |
| C | 12 | 34 | 9.5 | 337 | 12 | BG270991 | BG270991 1b08d10.Y |
| C | 13 | 34 | 9.5 | 855 | 17 | CNS01XEV | AL171616 Telraodon |
| C | 14 | 34 | 9.5 | 1078 | 12 | BF526014 | BF526014 602070191 |
| C | 15 | 33.8 | 9.5 | 550 | 14 | B0538836 | B0538836 MEST606-C |
| C | 16 | 33.8 | 9.5 | 634 | 13 | BM078579 | BM078579 MEST121-F |
| C | 17 | 33.4 | 9.4 | 604 | 13 | B1394070 | B1394070 pppin.pk0 |
| C | 18 | 33.4 | 9.4 | 634 | 13 | B1393243 | B1393243 pppin.pk0 |
| C | 19 | 33.2 | 9.3 | 259 | 9 | AA688556 | AA688556 vnl7b02.r |
| C | 20 | 33.2 | 9.3 | 457 | 9 | AA329374 | AA329374 EST33044 |
| C | 21 | 33.2 | 9.3 | 457 | 9 | AA672101 | AA672101 vll1407.r |
| C | 22 | 33.2 | 9.3 | 465 | 14 | M66777 | M66777 mel17b12.r1 |
| C | 23 | 33.2 | 9.3 | 486 | 9 | AA145845 | AA145845 mg35c08.r |
| C | 24 | 33.2 | 9.3 | 555 | 9 | AA238467 | AA238467 mx83c08.r |
| C | 25 | 33 | 9.2 | 363 | 13 | B1497901 | B1497901 sag29b11. |
| C | 26 | 33 | 9.2 | 422 | 9 | AA724936 | AA724936 a106c06.s |
| C | 27 | 33 | 9.2 | 544 | 13 | B1497969 | B1497969 sag30d01. |
| C | 28 | 33 | 9.2 | 556 | 14 | B0741626 | B0741626 sag20h07. |
| C | 29 | 33 | 9.2 | 583 | 17 | AO947804 | AO947804 Sheared D |
| C | 30 | 32.8 | 9.2 | 400 | 14 | BQ620730 | BQ620730 Talr1131D |
| C | 31 | 32.8 | 9.2 | 640 | 13 | BM604645 | BM604645 170006870 |
| C | 32 | 32.8 | 9.2 | 890 | 9 | AL568220 | AL568220 AL568220 |
| C | 33 | 32.6 | 9.1 | 521 | 17 | TA259E12Q | AL48386 T. brucei |
| C | 34 | 32.6 | 9.1 | 948 | 13 | B195658 | B195658 HVSME000 |
| C | 35 | 32.4 | 9.1 | 248 | 9 | AA369256 | AA369256 EST80663 |
| C | 36 | 32.4 | 9.1 | 381 | 10 | BE483949 | BE483949 170436 BA |
| C | 37 | 32.4 | 9.1 | 420 | 14 | BQ363954 | BQ363954 RC4-ST027 |
| C | 38 | 32.4 | 9.1 | 532 | 12 | BF230244 | BF230244 251749 BA |
| C | 39 | 32.4 | 9.1 | 536 | 14 | BQ315599 | BQ315599 CM0-CT001 |
| C | 40 | 32.4 | 9.1 | 546 | 12 | BG149197 | BG149197 naa24d10. |
| C | 41 | 32.4 | 9.1 | 547 | 14 | BQ373286 | BQ373286 MR1-F7020 |
| C | 42 | 32.4 | 9.1 | 832 | 9 | AL579156 | AL579156 AL579156 |
| C | 43 | 32.4 | 9.1 | 920 | 17 | CNS0062R | AL061710 Drosophila |
| C | 44 | 32.4 | 9.1 | 952 | 9 | AL564543 | AL564543 AL564543 |
| C | 45 | 32.2 | 9.0 | 574 | 12 | BE903935 | BE903935 601494720 |

ALIGNMENTS

RESULT 1
AY105813
LOCUS 1094 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PC0139085 mRNA sequence.
ACCESSION AY105813
VERSION AY105813.1 GI:21208891
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS Hainey/C.F., Dolan/M., Miao/G.H., Vogel/J.M., Whitsitt/M.S.,
Arthur/L.W., Hanafey/M., Morgante/M. and Tingey/S.V.
TITLE Maize Mapping Project/Dupont Consensus Sequences for Design of
Overgo Probes

JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1094)
AUTHORS Coe/E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
FEATURES
source location/Qualifiers
1..1094
/organism="Zea mays"
/db_xref="MaizeDB:638167"
/db_xref="taxon:4577"
/clone="PC0139085"


```
Query Match      8.7%; Score 31.2; DB 9; Length 624;
Best Local Similarity 50.3%; Pred. No. 0.47;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 140 TTTTAAATACCAATAAGATTGAGATTTCCGCCGGTAACCTTGTGACGNTGAGCAATG 199
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 401 TTTTCTTTAATACATCATCTAATGTGTTTACTTCTTACGTTTAACATCATTAATCAATG 342
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 TTCCGACGATGATCTCCAGCTTGCGCATGTGTTGATGCATGATGCGCCATTTCT 259
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 341 TTATCAACGATGATCTTTTGGCTTCTTAAGCTGTGTCAGACTTCTTAATGGCTTCT 282
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 CCTCTCTGACCACTTCGCGCGCTGCTC 288
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 TCTTTTCAAAATTCCTTTAGACGATCATC 253
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-925-637-51/c
; Sequence 51, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 51
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-51

Query Match      8.7%; Score 31.2; DB 10; Length 624;
Best Local Similarity 50.3%; Pred. No. 0.47;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 140 TTTTAAATACCAATAAGATTGAGATTTCCGCCGGTAACCTTGTGACGNTGAGCAATG 199
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 401 TTTTCTTTAATACATCATCTAATGTGTTTACTTCTTACGTTTAACATCATTAATCAATG 342
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 TTCCGACGATGATCTCCAGCTTGCGCATGTGTTGATGCATGATGCGCCATTTCT 259
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 341 TTATCAACGATGATCTTTTGGCTTCTTAAGCTGTGTCAGACTTCTTAATGGCTTCT 282
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 CCTCTCTGACCACTTCGCGCGCTGCTC 288
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 TCTTTTCAAAATTCCTTTAGACGATCATC 253
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-08-781-986A-152/c
; Sequence 152, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 152:
SEQUENCE CHARACTERISTICS:
LENGTH: 7953 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: Linear
US-08-781-986A-152

Query Match      8.7%; Score 31.2; DB 7; Length 7953;
Best Local Similarity 50.3%; Pred. No. 1.6;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 140 TTTTAAATACCAATAAGATTGAGATTTCCGCCGGTAACCTTGTGACGNTGAGCAATG 199
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 946 TTTTCTTTAATACATCATCTAATGTGTTTACTTCTTACGTTTAACATCATTAATCAATG 887
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 TTCCGACGATGATCTCCAGCTTGCGCATGTGTTGATGCATGATGCGCCATTTCT 259
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 886 TTATCAACGATGATCTTTTGGCTTCTTAAGCTGTGTCAGACTTCTTAATGGCTTCT 827
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 CCTCTCTGACCACTTCGCGCGCTGCTC 288
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 826 TCTTTTCAAAATTCCTTTAGACGATCATC 798
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-815-343-1076/c
; Sequence 1076, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon F.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1076
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(433)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-1076
```


RESULT 5
US-09-785-921A-3
; Sequence 3, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 3
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Ricinus communis
US-09-785-921A-3

Query Match 14.6%; Score 52.2; DB 10; Length 1731;
Best Local Similarity 49.6%; Pred. No. 2.1e-08;
Matches 129; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATCCTGTCGAGGCTTCGCCGCAATTAACAGTGGGAGTCNTCT 60
DB 640 GCAGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 699
QY 61 CCACCAACATGTACATGCTGAGTGGAGGCTTCGCCGCAATTAACAGTGGGAGTC 120
DB 700 GCACGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 759
QY 121 CAGCAGTCCAGATGATGATTTTATACCAATTAAGATTCGATTCGCCGCTAC 180
DB 760 CAAGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 819
QY 181 TTGTGACGNTGAGCAATGTTGCGCAGTGTATTCACCTTGGGAGTGTTCGAA 240
DB 820 AAATTCAGTGTGATGAGTGTATTAATCCCTATCATAGCTCTCATGTTATAGA 879
QY 241 TGCAGTGTGCGGCAATTCCTC 260
DB 880 TCGCAGCTTCACCATCTCTC 899

RESULT 6
US-09-785-921A-10
; Sequence 10, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 10
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: Ricinus communis
US-09-785-921A-10

Query Match 14.6%; Score 52.2; DB 10; Length 1805;
Best Local Similarity 49.6%; Pred. No. 2.2e-08;
Matches 129; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATCCTGTCGAGGCTTCGCCGCAATTAACAGTGGGAGTCNTCT 60
DB 648 GCAGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 707

QY 61 CCACCAACATGTACATGCTGAGTGGAGAGAGTGGGTCGACAAATCCACCAATCTC 120
DB 708 GCACGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 767
QY 121 CAGCAGTCCAGATGATGATTTTATACCAATTAAGATTCGATTCGCCGCTAC 180
DB 768 CAAGAGTTCATTAATGAGGAGAAATGCGCAGAGATTAGGTACACCGAGATCT 827
QY 181 TTGTGACGNTGAGCAATGTTGCGCAGTGTATTCACCTTGGGAGTGTTCGAA 240
DB 828 AAATTCAGTGTGATGAGTGTATTAATCCCTATCATAGCTCTCATGTTATAGA 887
QY 241 TGCAGTGTGCGGCAATTCCTC 260
DB 888 TCGCAGCTTCACCATCTCTC 907

RESULT 7
US-09-815-242-7727
; Sequence 7727, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7727
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1977)
US-09-815-242-7727

Query Match 9.1%; Score 32.6; DB 10; Length 1977;
Best Local Similarity 48.1%; Pred. No. 0.25;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 170 CCGCGGTAACCTTGTGACGNTGAGCAATGTTGCGAGTGTATTCACCTTGGGAGTCA 229
DB 881 CTTTCGTAATCTCTCCGAGTGCACCGACCTTCGCCGCAAGGCAATGTTGGCTACA 940
QY 230 TGTGTGCAATGCAATGTTGCGGCAATTCCTCTCTGACCACTTCGCGCTGCTCC 289
DB 941 AGTGTGATGCTGTGCGGCAAGGCGCGGATTCCTCTGCGCTTCACCGTGGCTGCAC 1000
QY 290 TAAGTCCGTGTGATGAGGCGCAACGATGTACCTGTGACTTTTCGAACCCACCTGTC 349

| | | | | | | | | |
|----|-----|----------|------------|----------|-------------|-----------|---------------|-----|
| QY | 181 | TTTTCAGC | GNNAGCAATG | TTGGAGAC | GTATCTCCAGC | TTGGGAA | CAATCGTTGCGAA | 240 |
| | | | | | | | | |
| Db | 679 | TTCTCAGC | TTGACCAATG | TTCCGACG | TCATGCCACG | CTTGGCGAT | CAATCTGTTGTA | 758 |
| | | | | | | | | |
| QY | 241 | TGCAGT | GTGTCGCCAT | 256 | | | | |
| | | | | | | | | |
| Db | 739 | TGCGAG | AGCGGCCAT | 754 | | | | |
| | | | | | | | | |

RESULT 2

Sequence 1 Application US/09347064A
 Patent No. US20020045208A1
 GENERAL INFORMATION:
 APPLICANT: Eck, Jürgen
 APPLICANT: Schmidt, Arno
 APPLICANT: Zinke, Holger
 TITLE OF INVENTION: Recombinant Fusion Proteins Based on
 TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
 TITLE OF INVENTION: album
 FILE REFERENCE: 09282-5
 CURRENT APPLICATION NUMBER: US/09/347,064A
 CURRENT FILING DATE: 1999-07-02
 EARLIER APPLICATION NUMBER: PCT/EP98/00009
 EARLIER FILING DATE: 1998-01-02
 EARLIER APPLICATION NUMBER: EP 97 10 0012.0
 EARLIER FILING DATE: 1997-01-02
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 762
 TYPE: DNA
 ORGANISM: Viscum album
 OS-09-347-064-1

| | | | | |
|---------------------------|--------|--------------------|-----------|-------------|
| Query Match | 51.0%; | Score 182.2; | DB 10; | Length 762; |
| Best Local Similarity | 81.9%; | Pred. NO. 2.4e-55; | | |
| Matches 208; Conservative | 0; | Mismatches 46; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|---|-----|
| Qy | 1 | GCAGATTCAATCCCATCUNGTGGAGGCTTCGCCGGAATTAACAGTGGAGATCMTCT | 60 |
| Db | 505 | GCCAGATTCAATCCCATCTTTATGGAGGGCTGCCCAATACATTAACAGTGGAGCTCATTTT | 564 |
| Qy | 61 | CCACCAAAATGATGATGCTGAGCTGGAGACGATTTGGGGTGCACATCCACCAAGTC | 120 |
| Db | 585 | CTCCGACAGCGTATCATGCTGAGACTGGAGACGATTTGGGGCCAAACATCCACCAAGTC | 624 |
| Qy | 121 | CAGCAGTCCAGGATGGCATTTTAAATACCAAAATTAAGATTGCGAGATTTCCGCGGTAAC | 180 |
| Db | 625 | CAGCATTTCAACCGATGGCGTTTAAATPAACCAATTCGGTTGGCTAATACCCCCCGGTAAC | 684 |
| Qy | 181 | TTTGTGACGNTGACCAATGTTCCGAGAGTGATCTCAACGCTTGGCATCATCTTTGTGAA | 240 |
| Db | 685 | TTTGTGACGTTGACCAATGTTCCGAGAGTGATCTCCAGACTTGGGATCATCTTTGTGTA | 744 |
| Qy | 241 | TGCAATGGTCCGCC | 254 |
| Db | 745 | TGCGGAGAGCGCCC | 758 |

RESULT 3

```

: Sequence 1, Application US/09785921A
: Patent No. US2002009434A1
: GENERAL INFORMATION:
: APPLICANT: Keener, William K.
: APPLICANT: Ward, Thomas E.
: TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
: FILE REFERENCE: LIT-PI-529
: CURRENT APPLICATION NUMBER: US/09/785,921A
: CURRENT FILING DATE: 2001-02-15
: NUMBER OF SEQ ID NOS: 17

```

```

; SEQ ID NO 1
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Ricinus communis
US-09-785-921A-1

```

| | | | | |
|-----------------------|-----------------|--------------------|-----------|--------------|
| Query Match | 15.1%; | Score 53.8; | DB 10; | Length 1698; |
| Best Local Similarity | 50.0%; | Pred. No. 5.6e-09; | | |
| Matches 130; | Conservative 0; | Mismatches 130; | Indels 0; | Gaps 0; |

1

[illegible]

RESULT 4

```

: Sequence 2, Application US/09785921A
:
: Patent NO. US2002009434A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Ward, Thomas E.
:
: TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
:
: TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
:
: FILE REFERENCE: LIT-PI-529
:
: CURRENT APPLICATION NUMBER: US/09/785,921A
:
: CURRENT FILING DATE: 2001-02-15
:
: NUMBER OF SEQ ID NOS: 17
:
: SEQ ID NO 2
:
: LENGTH: 1731
:
: TYPE: DNA
:
: ORGANISM: Ricinus communis
:
: US-09-785-921A-2

```

| | | | | |
|---------------------------|--------|--------------------|-----------|--------------|
| Query Match | 15.18; | Score 53.8; | DB 10; | Length 1731; |
| Best Local Similarity | 50.08; | Pred. No. 5.6e-09; | | |
| Matches 130; Conservative | 0; | Mismatches 130; | Indels 0; | Gaps 0 |

| | | | |
|----|-----|---|-----|
| QY | 1 | GCAGATTCATCCCATCTNTGTGGAGCTTGGCCGGCCAAATTAACTAGGGGAGTCTCT | 60 |
| | | | |
| Db | 640 | GCMAATTTTCCATATATTATGAGGGAGAAATTTGGCCACGAGAAATTAGTACCAACCGGAGATCT | 699 |
| | | | |
| QY | 61 | CCACCAAAACATGTATCATGCTGCTGAGCTGGAGACAGAGTTGGGGTCCACAATCCACCAAGTC | 120 |
| | | | |
| Db | 700 | GCACCAAGATCCTAGGCTAATTACACTGTGAAATGTTGGGGAGACATTTCCACTGCAATT | 759 |
| | | | |
| QY | 121 | CAGCAGTCCAMAGATGSCATTTTAAATACCAATTAAGATTGCGAGATTTCCGCCGGTAAAC | 180 |
| | | | |
| Db | 760 | CAGAAGTCTAACCAAGGAGCCTTTGTGTCCAAATTCACAGCAAAACAGTAATGATTTCC | 819 |
| | | | |
| QY | 181 | TTTGGAGGNTGAGCAATGTTTCGGACGCTGATATTCACACTGGGAGATCATGTTTGGAA | 240 |
| | | | |
| Db | 820 | AAATTCACGTGTACGATGTGAGTATTAATTCCTATCATAGCTCATGTGTGTATAGA | 879 |
| | | | |
| QY | 241 | TGCAGTGGTGGCCATTTCTC | 260 |
| | | | |
| Db | 880 | TTCGCACCTCCACCATCTGCT | 899 |
| | | | |

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 15:20:00 ; Search time 869.178 Seconds
(without alignments)
11953.474 Million cell updates/sec

Title: us-09-627-165e-13

Sequence: 1 gccacagatcaatccatcctt.....aacccacgtgcgcacgtcga 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl: *
8: gb_pr: *
9: gb_ro: *
10: gb_sts: *
11: gb_sy: *
12: gb_un: *
13: gb_vi: *
14: em_ba: *
15: em_fun: *
16: em_hum: *
17: em_in: *
18: em_mu: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_un: *
28: em_vi: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rtd: *
36: em_htg_mam: *
37: em_htg_vrl: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 353 | 98.9 | 357 | 6 | AX139583 | AX139583 Sequence |
| 2 | 220 | 64.4 | 522 | 6 | AX139585 | AX139585 Sequence |
| 3 | 225.8 | 63.2 | 1873 | 8 | AF369961 | AF369961 Viscum al |
| 4 | 218 | 61.1 | 1596 | 6 | AX019438 | AX019438 Sequence |
| 5 | 217.6 | 61.0 | 1598 | 6 | AX019435 | AX019435 Sequence |
| 6 | 216.4 | 60.6 | 1923 | 6 | AS68957 | AS68957 Sequence 1 |
| 7 | 216.4 | 60.6 | 1923 | 6 | AR164305 | AR164305 Sequence |
| 8 | 214.8 | 60.2 | 1602 | 6 | AY081149 | AY081149 Viscum al |
| 9 | 190.2 | 53.3 | 763 | 6 | AX019436 | AX019436 Sequence |
| 10 | 189 | 52.9 | 762 | 6 | AX019439 | AX019439 Sequence |
| 11 | 186.4 | 52.2 | 747 | 8 | AY081148 | AY081148 Viscum al |
| 12 | 185.8 | 52.0 | 705 | 6 | AR164307 | AR164307 Sequence |
| 13 | 185.8 | 52.0 | 756 | 6 | A90900 | A90900 Sequence 7 |
| 14 | 185.8 | 52.0 | 774 | 6 | AR164303 | AR164303 Sequence |
| 15 | 184.2 | 51.6 | 768 | 6 | AX139575 | AX139575 Sequence |
| 16 | 182.6 | 51.1 | 762 | 6 | AX139571 | AX139571 Sequence |
| 17 | 182.6 | 51.1 | 762 | 8 | AF508914 | AF508914 Viscum al |
| 18 | 182.2 | 51.0 | 762 | 6 | A90894 | A90894 Sequence 1 |
| 19 | 182.2 | 51.0 | 762 | 6 | AX139573 | AX139573 Sequence |
| 20 | 179.4 | 50.3 | 768 | 6 | AX019440 | AX019440 Sequence |
| 21 | 176.8 | 49.5 | 753 | 8 | AF508916 | AF508916 Viscum al |
| 22 | 172.4 | 48.3 | 747 | 8 | AF508915 | AF508915 Viscum al |
| 23 | 90 | 25.2 | 1596 | 6 | AX019441 | AX019441 Sequence |
| 24 | 78.6 | 22.0 | 762 | 6 | AX019442 | AX019442 Sequence |
| 25 | 78.6 | 22.0 | 768 | 6 | AX019443 | AX019443 Sequence |
| 26 | 55.4 | 15.5 | 1947 | 6 | E01355 | E01355 cDNA encodl |
| 27 | 55.2 | 15.5 | 1825 | 6 | AX108095 | AX108095 Sequence |
| 28 | 54.2 | 15.2 | 1837 | 6 | AX108130 | AX108130 Sequence |
| 29 | 54.2 | 15.2 | 1843 | 6 | AX108123 | AX108123 Sequence |
| 30 | 54.2 | 15.2 | 1849 | 6 | AX108116 | AX108116 Sequence |
| 31 | 53.8 | 15.1 | 839 | 6 | I07422 | I07422 Sequence 34 |
| 32 | 53.8 | 15.1 | 839 | 6 | I08354 | I08354 Sequence 10 |
| 33 | 53.8 | 15.1 | 873 | 6 | A04860 | A04860 Artificial |
| 34 | 53.8 | 15.1 | 873 | 6 | A04867 | A04867 Artificial |
| 35 | 53.8 | 15.1 | 1050 | 6 | I02107 | I02107 Sequence 3 |
| 36 | 53.8 | 15.1 | 1050 | 6 | I07418 | I07418 Sequence 30 |
| 37 | 53.8 | 15.1 | 1050 | 6 | I08353 | I08353 Sequence 9 |
| 38 | 53.8 | 15.1 | 1140 | 6 | A23903 | A23903 partial seq |
| 39 | 53.8 | 15.1 | 1140 | 6 | A23989 | A23989 pIC1 1102 p |
| 40 | 53.8 | 15.1 | 1140 | 6 | AR059851 | AR059851 Sequence |
| 41 | 53.8 | 15.1 | 1140 | 6 | I25094 | I25094 Sequence 15 |
| 42 | 53.8 | 15.1 | 1689 | 8 | S40368 | S40368 agglutinin |
| 43 | 53.8 | 15.1 | 1695 | 6 | A04859 | A04859 Artificial |
| 44 | 53.8 | 15.1 | 1695 | 6 | A10437 | A10437 Synthetic n |
| 45 | 53.8 | 15.1 | 1695 | 6 | E00828 | E00828 DNA encodln |

ALIGNMENTS

RESULT 1
AX139583
LOCUS AX139583 357 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 13 from Patent EP1074560.
ACCESSION AX139583
VERSION AX139583.1 GI:14275213
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
REFERENCE 1 (bases 1 to 357)
Kim,U., Song,S., Suh,B., Lee,K., Doo,M., Kwak,J., Song,B., Yoon,T.,
Kang,T. and Park,C.

Pred. No. is the number of results predicted by chance to have a

TITLE Crude extract from viscum album coloratum, and proteins and lectins isolated therefrom
JOURNAL Patent: EP 1074560-A 13 07-FEB-2001;
FEATURES Mistle Biotech Co., Ltd. (KR)
SOURCE Location/Qualifiers
 1. 357
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 <1..>357
CDS /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC39972.1"
 /db_xref="GI:14275214"
 /translation="ARENPIMRLRQINSSESPNNMILETSWROSTOVQOSK
 DGINTOIRLQISAGNEFTYSNVRDISSLAIMLFECSGREFSSLDHPSPULLNSYVD
 AANDVTCXSEPTVRIY"
BASE COUNT 77 a 103 c 87 g 86 t 4 others
ORIGIN
 Query Match 98.9%; Score 353; DB 6; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.5e-98;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0Y 1 GCCAATTCATCCCATGNTGAGGCTTGCCGCGCAATTACAGTGGGAGTCNTCT 60
 DB 1 GCCAATTCATCCCATGNTGAGGCTTGCCGCGCAATTACAGTGGGAGTCNTCT 60
 0Y 61 CCACCAACATGTACATGCTCGAGCTGGAGAGAGATTGGGTCGACATCCACCAATGC 120
 DB 61 CCACCAACATGTACATGCTCGAGCTGGAGAGAGATTGGGTCGACATCCACCAATGC 120
 0Y 121 CAGCAGTCCAGAGATGGAATTTTAAATCCCAATTAATTCAGATTTCCGCGGTAAAC 180
 DB 121 CAGCAGTCCAGAGATGGAATTTTAAATCCCAATTAATTCAGATTTCCGCGGTAAAC 180
 0Y 181 TTTGTGAGGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGGATCATGTTGTGAA 240
 DB 181 TTTGTGAGGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGGATCATGTTGTGAA 240
 0Y 241 TGCATGTGTCGGCATTTCTCTCTGTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 DB 241 TGCATGTGTCGGCATTTCTCTCTGTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 0Y 301 GTGATGGCGGCAAGATGTACCGTACCTTTTCCGACACCCGCGTGCATGCTGA 357
 DB 301 GTGATGGCGGCAAGATGTACCGTACCTTTTCCGACACCCGCGTGCATGCTGA 357
RESULT 2
AX139585 522 bp DNA linear PAT 30-MAY-2001
LOCUS AX139585
DEFINITION Sequence 15 from Patent EP1074560.
ACCESSION AX139585
VERSION AX139585.1 GI:14275215
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
AUTHORS Kim, J., Song, S., Suh, B., Lee, K., Do, M., Kwak, J., Song, B., Yoon, T.,
 Kang, T., and Park, C.
TITLE Crude extract from viscum album coloratum, and proteins and lectins
 isolated therefrom
JOURNAL Patent: EP 1074560-A 15 07-FEB-2001;
FEATURES Mistle Biotech Co., Ltd. (KR)
SOURCE Location/Qualifiers
 1. 522
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 <1..>522
CDS /note="unnamed protein product"

/codon_start=1
 /protein_id="CAC39973.1"
 /db_xref="GI:14275216"
 /translation="YNDLERYAGHRDPIGLIEELIOSVALRYPGSGTAAQASLII
 LIQMISEARFNPIDFWARQYINSSEFPLDMIMLETSWROSTOVQOSDGVENN
 PFRIGISTGNEFTLSNVRDISSLAIMLFVCHDRBSSDVRWPLIVIRVLENSGAVD
 DVNCTASEPTVRIY"
BASE COUNT 112 a 149 c 132 g 129 t
ORIGIN
 Query Match 64.4%; Score 230; DB 6; Length 522;
 Best Local Similarity 78.7%; Pred. No. 2.7e-60;
 Matches 288; Conservative 0; Mismatches 69; Indels 9; Gaps 1;
 0Y 1 GCCAATTCATCCCATGNTGAGGCTTGCCGCGCAATTACAGTGGGAGTCNTCT 60
 DB 157 GCGAATTCATCCCATGNTGAGGCTTGCCGCGCAATTACAGTGGGAGTCNTCT 60
 0Y 61 CCACCAACATGTACATGCTCGAGCTGGAGAGAGATTGGGTCGACATCCACCAATGC 120
 DB 217 CTTCGCCACATGTACATGCTCGAGCTGGAGAGAGATTGGGTCGACATCCACCAATGC 120
 0Y 121 CAGCAGTCCAGAGATGGAATTTTAAATCCCAATTAATTCAGATTTCCGCGGTAAAC 180
 DB 277 CAGCAGTCCAGAGATGGAATTTTAAATCCCAATTAATTCAGATTTCCGCGGTAAAC 180
 0Y 181 TTTGTGAGGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGGATCATGTTGTGAA 240
 DB 337 TTTGTGAGGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGGATCATGTTGTGAA 240
 0Y 241 TGCATGTGTCGGCATTTCTCTCTGTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 DB 397 TGCATGTGTCGGCATTTCTCTCTGTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 0Y 301 GTGGA-----TGGGCGCAACGATGTACCTGACCTTTTCCGACACCCGCGTGC 351
 DB 457 TGGGAAGAAAGCGGCGCGTGCAGATGTACCTGACCTGACCTTCCGACACCCGCGTGC 351
 0Y 352 ATCGTA 357
 DB 517 ATCGTA 522
RESULT 3
AF369961 1873 bp DNA linear PLN 12-DEC-2001
LOCUS AF369961
DEFINITION Viscum album subsp. coloratum VCA precursor, gene, complete cds.
ACCESSION AF369961
VERSION AF369961.1 GI:17529700
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
AUTHORS Park, W.-B. and Lyu, S.
TITLE Cloning of Viscum album subsp. coloratum (Korean mistletoe)
JOURNAL Biochem. Biophys. Res. Commun. (2002) In press
REFERENCE 2 (bases 1 to 1873)
AUTHORS Park, W.-B. and Lyu, S.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2001) Chemistry, Seoul Women's University, 126,
 Kongnung-2dong, Nowon-gu, Seoul 139-774, Korea
FEATURES Location/Qualifiers
 1. 1873
 /organism="Viscum album subsp. coloratum"
 /sub_species="coloratum"
 /db_xref="taxon:159976"
 <138..>1835
 /product="VCA precursor"
 138..1835
 /codon_start=1

| | | | | |
|------------|--|-------|-------|-------|
| BASE COUNT | 456 a | 458 c | 492 g | 467 t |
| ORIGIN | /product="VCA precursor"
/protein_id="AAL0417.1"
/db_xref="GI:17529701"
/translation="MNALASRAWVSCFLMLGLVFGATYKAETKRYERYLRLLRVTHQ
TTGGEYFRFTLLRLDYVSSGSNEIPLRLRSTIPVDAORFVLELTNGGDSITAAI
IDWYLVVAYVAGDQDSYFLRPADPAGAEHRLHFTSTSLPFGTSYDLELYAGHROO
IPCIPELLIVSYALDRYSGGSTRPAORSLTLLIOMISEARFNPIMPARYISNGSES
FLDMYHLELETSWGGQSTVOYQDSTDCVFNPNFPLGISTGFTVLSNRYIASLAIH
LFVCDRPPSSDVAYWPLVLRPVLENSGADVDTCTAASEPVRIVRGHGLCDVRDEK
FYENPRLQLPMDPNQMLTIRRDSTYVAGSGNCLTYYTAAVYVIFDCNPAVEATPL
MQLNGNGLTIINPSNLVIGAASSGSGTLLIVQTVYRGLGOMLAGNTPAPEYVITYGF
GILCMENGAASIVETCGSGKENQKMLYGGSIIRPOMONQDCLTSGDGSIVSENIY
SCSAGSGSGRHEFNECTIILNLNGILVMDVAQSNPSLRITIIYATPKRPNMLPVP
138..203
204..956
/product="VCA alpha chain"
/product..1061
957..1061
/note="Region: linker peptide"
1062..1832
/product="VCA beta chain"
1833..203 | | | |

| | | | | |
|---------------------------|-------|--------------------|-----------|--------------|
| Query Match | 63.2% | Score 225.8; | DB 8; | Length 1873; |
| Best Local Similarly | 78.1% | Pred. No. 6.5e-59; | | |
| Matches 285; Conservative | 0; | Mismatches 71; | Indels 9; | Gaps 1; |

QY 1 GCGAGATTCATCCCATCTGTGAGGGCTGCGCCGGCAATTTAAAGTGGGAGTCNCT 60
 Db 735 GCGAGATTCATCCCATCTTTTGGAGGGCTGCCCATACATTTAAAGCGGGGAGTCATTT 794
 QY 61 CCACCAACATGTACATGCTGAGCTGGAGACGATTTGGGCTCGACAAATCCACCAAGTC 120

OY 121 CAGCAGTCCAMGANGCATTTTAAATACCCAATAAAGATTGCAGATTTCCGCCCGTAAC 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 855 CAGCAGTCTACGAGTAGGGTTTTTAATAACCATTTCGGTGGAATATCCACCGGTAAAC 914

OY 181 TTTTGACAGANTAGCAATGTTGCCACAGCTGANCTCAGCTTGGCGATCATGTTGTTGCANA 240

Db 915 TTCGAGCTTGAGCAATGTTGGCAGCGGATCGGCACCTTAGCCATCACTGTTGTTTGA 974

Oy 241 TGCAGTGTGGCCATTCCTCTCTCGACCAACCTTCGCGCGCTGCTCTCAAGTCCGTC 300

Db 975 TGTAGGAGCCGACACATCTTCCCTCGACGCGGCTATTGGCGCGTGTCTATAGACCCGTC 1034

Db 1035 TTGGAAATAGCGGCGCCCTCGACAGCTGTACTCGATCTTCGACACCCACCGTGGG 1094

Qy 352 ATCGT 356

Db 1095 ATTGT 1099

| | | | | |
|------------|-----------------------------------|-------------|-----|-----------------|
| RESULT 4 | | | | |
| LOCUS | AX019438 | 1586 bp | DNA | linear |
| DEFINITION | Sequence 15 from Patent WO940109. | | | |
| ACCESSION | AX019438 | | | |
| VERSION | AX019438.1 | GI:10043377 | | PAT 07-SEP-2000 |

| SOURCE | synthetic construct. |
|-----------|---|
| ORGANISM | synthetic construct. |
| REFERENCE | artificial sequences. |
| AUTHORS | 1 (bases 1 to 156) |
| TITLE | Walters, P., Stiefel, T., Voelter, W. and Morris, P. |
| JOURNAL | Recombinant mastlctoe lectins |
| | Patent: WO 9940109-A 15 12-AUG-1999; |
| | WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH |
| FEATURES | (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) |
| | Location/Qualifiers |

| | |
|------------|---------------------------------|
| SOURCE | 1. 1596 |
| | /organism="synthetic construct" |
| | /db_xref="taxon:32630" |
| BASE COUNT | 387 a 418 c 419 g 372 t |
| ORIGIN | |

| | | | | |
|-----------------------|--------------|--------------------|----------------|--------------|
| Query Match | 61.1%; | Score 218; | DB 6; | Length 1596; |
| Best Local Similarity | 77.5%; | Pred. No. 1.7e-56; | | |
| Matches 276; | Conservative | 0; | Mismatches 74; | Indels 6; |
| | | | | Gaps 1; |

[illegible]

QY 121 CAGCAGTCCAGGATGGCATTTTAAATACCCAAATAGATTCAGATTTCCSCCGGTAC 180
||||| | | ||||| ||||| | | ||| | | ||||| |
Db 619 CAGCATTCACCGATGGCCGTTTTTAATAACCCAAATTCGGTTGAGCTATACCCCCCGGTAC 678
QY 181 TTTGTGACANTGAGCAATCTCGCGACGATCTCAGCTTGCGATCATGTTGTTGCA 240

[illegible]

| Db | 799 | AT | -----AGCCGATCATGTTCACCTGCAGTCGTCGGAACCTACGGTGGCGGATGT | 848 |
|----------|----------|---------|---|-----------------|
| RESULT 5 | | | | |
| AX019435 | | | | |
| 100% | | | | |
| | AX019435 | 1598 bp | DNA | linear |
| | | | | PAT 07-spp-2000 |

| | |
|------------|------------------------------------|
| DEFINITION | Sequence 12 from Patent WO9940109. |
| ACCESSION | AX019435 |
| VERSION | AX019435.1 |
| KEYWORDS | GI:10043374 |
| SOURCE | synthetic construct. |
| ORGANISM | synthetic construct |
| | artificial sequences |

| FEATURES | REFERENCE |
|----------|---|
| AUTHORS | 1 (bases 1 to 1598) |
| TITLE | Welters, P., Stiefel, T., Voelter, W. and Morris, P. |
| JOURNAL | Recombinant mistletoe lectins |
| | Patent: WO 99/0109-A 12 12-AUG-1999; |
| | WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH |
| | (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) |
| | Location/Qualifiers |

| | | |
|--------------|---------|---------------------------------|
| source | 1. 1598 | /organism="synthetic construct" |
| | | /db_xref="taxon:32630" |
| misc_feature | 319 | |
| misc_feature | 1322 | |
| BASE COUNT | 373 a | 396 c 392 g 356 t 81 others |
| ORIGIN | | |

| Query Match | 61.0% | Score 217.6; | DB 6; | Length 1598; |
|-----------------------|------------------|--|-----------|--------------|
| Best Local Similarity | 74.7%; | Pred. No. 2.2e-56; | | |
| Matches 266; | Conservative 13; | Mismatches 71; | Indels 6; | Gaps 1 |
| QY | 1 | GCAGATTCAATCCATCAGTGTGGAGCCTCGCGGCAATTACAGTGGGAGTCNTCC | 60 | |

QY 61 CCACCAACATGTATCATGCTCGAGCTGGAGAGCATGTTGGGCTGCACAATCCACCAAGTC 120
| ||| ||:||||||| ||||||||| ||||||||| ||||||||| |||||
Db 560 CTGCACGACCTGTACATGCTGGAGCTGGAGACGACGATTTGGGGCCAAACAATCCACGCAAGTC 619
QY 121 CACACGTCCAAGATGGCATTTTAAATACCAAAATAAGATTTGCAGATTTCCGGCCGGTAAC 180

|||||
Db 620 CAGCATTTCAACCGATGGCTTTTATAAACCAMTTCGGTGGCTATATACVCMYCGTAC 679
QY 181 TTGTGAGATGAGCAATGTTCGACGATGATCCAGCTTGCGCATATGTTGTGCA 240
Db 680 TTTCGACGCTTMCYTAAGTTGCGKMYGTGATCGCATGTCGCGATCATGTTGTTTA 739
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTGCGCGCTGCTCTTAAGTCCGTC 300
Db 740 TCGGAGAGACGGGCATCTCTCTGACGCGCTATGTGGCCGCTGCTATACGACCCGTG 799
QY 301 GTGATGCGGCCCAAGATGTACCTGCACTNTTTCGAAACCACCGTGGCATGCT 356
Db 800 AT-----AGCCGATGATGTTACCGCATGCTTGGAAACCTACCGTGGCATGCT 849

RESULT 6
A58957
LOCUS A58957 1923 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 1 from Patent WO9701636.
ACCESSION A58957
VERSION A58957.1 GI:3714428
KEYWORDS
SOURCE Viscum album.
ORGANISM Viscum album.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Santalales; Viscaceae; Viscum.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Lentzen, H., Eck, J., Baur, A. and Zinke, H.
TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: WO 9701636-A 16-JAN-1997;
MADAUS AG KOELN (DE)
FEATURES
source location/Qualifiers
1..1923
/organism="Viscum album"
/db_xref="taxon:3972"
/tissue_type="LEAF"
55..1749
/codon_start=1
/product="PREPROMISTLETOE-LECTIN"
/protein_id="CAA03513.1"
/db_xref="GI:3714428"
/translation="MNAVMDSRRAMASCFMLGLVFGATVKAETKESYERLRURVTHQ
TGDGEYRFTLLRDYVSSGFSNETPILRQSTIPVSDARFVLEVLINGGDSITAA
IDVNLVYVYAGDOSYFLDAPRGAEFTLFTGPRSSLPFGNSPDLERYGHDO
IPGLIDLIOSVTLALRPGSGTRQASLILILOMISEARFPLIMRAAOYNSGAS
FLPPLVMLEETSMGOSTOVHSTGCFENNRILAIPEENFTLVNRDYLSTLAI
LFNGERSSSEYRMYPLVIRPIADVTCSASEPVRIVGRNMCVDNRDDFRDGN
QIOLMPKSNNDPQMLTKRDGTIRNSGCLTYYGTAGVYVIMIDCNVAVEATLM
QIWMNGTIIINPRSNLVLAASSGIGKTLIVQTLDTLGGMLAGNDAPREVTYGER
DLMESNGSVWETCVSSQKNORMALYGGGSRPKONODGCLTGRDSVSTVNIYS
CSAGSGQRWVFNEGAILMLKNGLAMDVANQPKLRILIIYATKRPQMLPVP"
55..153
sig_peptide
mat_peptide 154..900
/product="A-CHAIN"
mat_peptide 901..957
/product="PROPEPTIDE"
mat_peptide 958..1746
/product="B-CHAIN"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN

Query Match 60.6%; Score 216.4; DB 6; Length 1923;
Best Local Similarity 77.2%; Pred. No. 5,4e-56;
Matches 275; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

QY 1 GCCAGATTCAATCCCATCCTGTGAGGCTTCGCCGCAAAATTAACAGTGGGAGTCTCT 60
Db 652 GCCAGATTCAATCCCATCTTATGAGGCTTCGCCCAATTAACAGTGGGAGTCTCT 711
QY 61 CCACCAAAACATGTACATGCTCGAGCTGGAGAGAGTTGGGGTCCGACAAATCCACCAAGTC 120
Db 712 CTGCCAGAGCGTGTACATGCTGAGAGCTGGAGAGAGTTGGGGCAACAATCCACGCAAGTC 771
QY 121 CACAGTCCCAAGATGGCATTTTAAATACCAATTAAGATTACAGATTTCGCCGCTAAC 180
Db 772 CAGCATTTCAACCGATGCGTGTAAATACCAATTCGGTGTGCTATACCCCGGTAAC 831
QY 181 TTGTGACGNTGAGCAATGTTCGAGAGTATCCGACGCTTGGCGATCATGTTGTGCA 240
Db 832 TTGCTGACGTTGACCAATGTTCCGAGAGTATGCCAGCTTGGCGATCATGTTGTTGA 891
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTGCGCGCTGCTCTTAAGTCCGTC 300
Db 892 TCGGAGAGCGGCCATCTCTCTGAGAGTGCCTTAATGGCGCGCTGTCATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTGCACTNTTTCGAAACCACCGTGGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCAAGTGTTCGGAACCTACCGTGGCATGCT 1001

|||||
Db 712 CTGCCAGAGCGTGTACATGCTGAGCTGGAGACGAGTTGGGGCAACAATCCACCAAGTC 771
QY 121 CAGAGTCCCAAGATGGCATTTTAAATACCAATTAAGATTGCAATTCGCCGCTAAC 180
Db 772 CAGCATTTCAACCGATGCGTGTAAATACCAATTCGGTGTGCTATACCCCGGTAAC 831
QY 181 TTGTGACGNTGAGCAATGTTCGAGAGTATCCGACGCTTGGCGATCATGTTGTGCA 240
Db 832 TTGCTGACGTTGACCAATGTTCCGAGAGTATGCCAGCTTGGCGATCATGTTGTTGA 891
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTGCGCGCTGCTCTTAAGTCCGTC 300
Db 892 TCGGAGAGCGGCCATCTCTCTGAGAGTGCCTTAATGGCGCGCTGTCATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTGCACTNTTTCGAAACCACCGTGGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCAAGTGTTCGGAACCTACCGTGGCATGCT 1001

RESULT 7
A164305
LOCUS A164305 1923 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 34 from patent US 6271368.
ACCESSION A164305
VERSION A164305.1 GI:16235409
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Lentzen, H., Eck, J., Baur, A. and Zinke, H.
TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: US 6271368-A 34-OCT-2001;
FEATURES
source location/Qualifiers
1..1923
/organism="unknown"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN

Query Match 60.6%; Score 216.4; DB 6; Length 1923;
Best Local Similarity 77.2%; Pred. No. 5,4e-56;
Matches 275; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

QY 1 GCCAGATTCAATCCCATCCTGTGAGGCTTCGCCGCAAAATTAACAGTGGGAGTCTCT 60
Db 652 GCCAGATTCAATCCCATCTTATGAGGCTTCGCCCAATTAACAGTGGGAGTCTCT 711
QY 61 CCACCAAAACATGTACATGCTCGAGCTGGAGAGAGTTGGGGTCCGACAAATCCACCAAGTC 120
Db 712 CTGCCAGAGCGTGTACATGCTGAGAGCTGGAGAGAGTTGGGGCAACAATCCACGCAAGTC 771
QY 121 CACAGTCCCAAGATGGCATTTTAAATACCAATTAAGATTACAGATTTCGCCGCTAAC 180
Db 772 CAGCATTTCAACCGATGCGTGTAAATACCAATTCGGTGTGCTATACCCCGGTAAC 831
QY 181 TTGTGACGNTGAGCAATGTTCGAGAGTATCCGACGCTTGGCGATCATGTTGTGCA 240
Db 832 TTGCTGACGTTGACCAATGTTCCGAGAGTATGCCAGCTTGGCGATCATGTTGTTGA 891
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTGCGCGCTGCTCTTAAGTCCGTC 300
Db 892 TCGGAGAGCGGCCATCTCTCTGAGAGTGCCTTAATGGCGCGCTGTCATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTGCACTNTTTCGAAACCACCGTGGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCAAGTGTTCGGAACCTACCGTGGCATGCT 1001

RESULT 8
AY081149
LOCUS AY081149 1602 bp mRNA linear PLN 17-MAR-2002
DEFINITION Viscum album lectin chain A isoform 1 precursor, mRNA, partial cds.
ACCESSION AY081149

| QY | 301 | GTGATGCGGCCCAACATGTCACCTGACTGTTTCCGAGCCACCGGTGGCATGCT | 356 |
|--|---|---|--------|
| | 1 | | |
| | 799 | AT-----AGCCGATGATGTTACCTGCAGTGTTCGGAACCTACGGTGGCGATTTGT | 848 |
| RESULT 9 | | | |
| LOCUS | AX019436 | 763 bp | DNA |
| DEFINITION | Sequence 13 from Patent WO940109. | | linear |
| ACCESSION | AX019436 | | |
| VERSION | AX019436.1 | GI:10043375 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | 1 (bases 1 to 763) | | |
| TITLE | Welters,P., Stiefel,T., Voelter,W. and Morris,P. | | |
| JOURNAL | Recombinant mistletoe lectines
Patent: WO 940109-A 13 12-AUG-1999;
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) | | |
| FEATURES | | | |
| source | 1..763 | | |
| | /organism="synthetic construct" | | |
| | /db_xref="taxon:32630" | | |
| misc_feature | 163 a 215 c 175 g 170 t 40 others | | |
| BASE COUNT | misc_feature 163 a 215 c 175 g 170 t 40 others | | |
| ORIGIN | | | |
| Query Match | 53.3%; Score 190.2; DB 6; Length 763; | | |
| Best Local Similarity | 78.4%; Pred. No. 6.6e-46; | | |
| Matches 207; Conservative 13; Mismatches 44; Indels 0; Gaps 0; | | | |
| QY | 1 | GCCAGATTCAAATCCATCTMTGTGGAGCGCTGCGCCGCAATTAACAGTGGGAGTCNTCT | 60 |
| | | | |
| Db | 500 | GCCAGATTCATATCCATCTTATGAGAGKMYGCCAAKAYATTAACAGTGGGAGTTCATTT | 559 |
| QY | 61 | CCACCAACATGTATCATGTGCTCGAGCTGAGAGCAGAGTGGGTCGACATCCACCAATGC | 120 |
| | | | |
| Db | 560 | CTGCGACAGACTGTACATGCTGAGCTGAGAGAGATTTGGGGCAATCCACCGCATGC | 619 |
| QY | 121 | CAGCAGTCCAGAGATGGCATTTTATATACCCAAATAAGATTGACATTTCCGCCGATAC | 180 |
| | | | |
| Db | 620 | CAGATTCACACCGATGGCGTTTAAATAACCAATMYCGTTGGCATATATCTMCGGTAC | 679 |
| QY | 181 | TTTGTGACGNTGACATGTTCCGAGCAGATCTCCAGCTTGGCGATCATGTTGCGAA | 240 |
| | | | |
| Db | 680 | TTTGTGACGTTGMCATATGTTGCCKMYTGATGCGCAGCTGGCATCATGTTGTTGTA | 739 |
| QY | 241 | TGCAGTGTGCGCCATTCCTCTCT | 264 |
| | | | |
| Db | 740 | TGCGAGAGCGGCCATCTCTCT | 763 |
| RESULT 10 | | | |
| LOCUS | AX019439 | 762 bp | DNA |
| DEFINITION | Sequence 16 from Patent WO940109. | | linear |
| ACCESSION | AX019439 | | |
| VERSION | AX019439.1 | GI:10043378 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | 1 (bases 1 to 762) | | |
| TITLE | Welters,P., Stiefel,T., Voelter,W. and Morris,P. | | |
| JOURNAL | Recombinant mistletoe lectines
Patent: WO 940109-A 16 12-AUG-1999;
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) | | |
| FEATURES | | | |
| source | 1..762 | | |
| | Location/Qualifiers | | |

DEFINITION Sequence 7 from Patent WO9829540.
 ACCESSION A90900
 VERSION A90900.1 GI:6739438
 KEYWORDS
 SOURCE European mistletoe.
 ORGANISM
 Viscum album
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 1 (bases 1 to 756)
 ECK, J. and Schmidt, A.
 RECOMBINANT FUSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS
 OF EUROPEAN MISTLETOE VISCUM ALBUM
 Patent: WO 9829540-A 7 09-JUL-1998;
 ECK JUERGEN (DE); SCHMIDT ARNO (DE)
 location/Qualifiers
 FEATURES
 source
 1..756
 /organism="Viscum album"
 /db_xref="taxon:3972"
 /tissue="leaf"
 1..756
 /gene="MLA"
 <1..>756
 /gene="MLA"
 /function="TOXIC COMPONENT"
 /note="FIG. 11.A: NUKLEOTIDSEQUENZ UND ABGELEITETE AMINO
 ACIDSEQUENZ VON RMLA"
 /codon_start=1
 /product="RECOMBINANT MISTLETOE LECTIN A-CHAIN."
 /protein_id="CAB69396.1"
 /db_xref="GI:6739439"
 /translation="VERIRLRVHQTGEYFRITLRLDYVSSGSFNSNELPIRROST
 IPVSDAORFVLELTNOGDSITPAIDVTNLVVAVAGDSIFLRAPKAEHLFT
 GTTSSLPFNQSYDLERYAGHDOIPIGLIDLIQSVTALEFPGSSTRGTQARSLILLI
 OMISEAARFNPIIMRROYINSASFLPDYMLELETSMWQOSTOYOVHSTDVGFNNDI
 RLAIIPGNEFVLTNRDVIASLAIMLFVCGRPS"
 BASE COUNT 167 a 229 c 181 g 179 t
 ORIGIN
 Query Match 52.0%; Score 185.8; DB 6; Length 756;
 Best Local Similarity 82.4%; Pred. No. 1.5e-46;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTCCCGGCAATTAACAGTGGAGTCNCT 60
 |||||||
 DB 499 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCAAATACAGTGGCGTCATT 558
 |||||||
 QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGGTCGACAAATCCACCAAGTC 120
 |||||
 DB 559 CTGCCAGAGCTGTACATGCTGAGAGAGAGAGAGTGGGCGCAACATCCACGCAAGTC 618
 |||||
 QY 121 CAGCAGTCCAGAGATGCAATTTTAAATACCAATTAAGTTCGAGATTTCCCGCGGTAA 180
 |||||
 DB 619 CAGCATTCAACGATGCGCTTTTAAATACCAATTCGCTGCTATACCCCCGGTAA 678
 |||||
 QY 181 TTGTGTACGTTGACATGTCGCGAGCGATCTCCAGTTGGCGATCTGTTGTGAA 240
 |||||
 DB 679 TTGTGTACGTTGACATGTCGCGAGCGATGCGCAGCTGCGCATCTGTTGTTGA 738
 |||||
 QY 241 TGCAGTGTGCGCCAT 256
 |||||
 DB 739 TGCAGAGAGCGGCCAT 754
 |||||
 RESULT 14
 ARI64303 774 bp DNA linear PART 17-OCT-2001
 LOCUS
 DEFINITION Sequence 30 from patent US 6271368.
 ACCESSION ARI64303
 VERSION ARI64303.1 GI:16235405
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

Unclassified.
 1 (bases 1 to 774)
 LENTZEN, H., ECK, J., BAUR, A. and ZINKE, H.
 RECOMBINANT MISTLETOE LECTIN (RML)
 Patent: US 6271368-A 30 07-AUG-2001;
 location/Qualifiers
 FEATURES
 source
 1..774
 /organism="unknown"
 BASE COUNT 173 a 232 c 185 g 184 t
 ORIGIN
 Query Match 52.0%; Score 185.8; DB 6; Length 774;
 Best Local Similarity 82.4%; Pred. No. 1.5e-46;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTCCCGGCAATTAACAGTGGAGTCNCT 60
 |||||||
 DB 505 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCAAATACAGTGGCGTCATT 564
 |||||||
 QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGGTCGACAAATCCACCAAGTC 120
 |||||
 DB 565 CTGCCAGAGCTGTACATGCTGAGAGAGAGAGAGTGGGCGCAACATCCACGCAAGTC 624
 |||||
 QY 121 CAGCAGTCCAGAGATGCAATTTTAAATACCAATTAAGTTCGAGATTTCCCGCGGTAA 180
 |||||
 DB 625 CAGCATTCAACCGATGCGCTTTTAAATACCAATTCGCTATACCCCCGGTAA 684
 |||||
 QY 181 TTGTGTACGTTGACATGTCGCGAGCGATCTCCAGTTGGCGATCTGTTGTGAA 240
 |||||
 DB 685 TTGTGTACGTTGACATGTCGCGAGCGATGCGCAGCTTGGCGATCTGTTGTGA 744
 |||||
 QY 241 TGCAGTGTGCGCCAT 256
 |||||
 DB 745 TGCAGAGAGCGGCCAT 760
 |||||
 RESULT 15
 AX139575 768 bp DNA linear PART 30-MAY-2001
 LOCUS
 DEFINITION Sequence 5 from Patent EP1074560.
 ACCESSION AX139575
 VERSION AX139575.1 GI:14275209
 KEYWORDS
 SOURCE Viscum album subsp. coloratum.
 ORGANISM Viscum album subsp. coloratum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 1 (bases 1 to 768)
 KIM, J., SONG, S., SUH, B., LEE, K., DOO, M., KWAK, J., SONG, B., YOON, T.,
 KANG, T. and PARK, C.
 Crude extract from Viscum album coloratum, and proteins and lectins
 isolated therefrom
 Patent: EP 1074560-A 5 07-FEB-2001;
 Mistle Biotech Co., Ltd. (KR)
 location/Qualifiers
 FEATURES
 source
 1..768
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 BASE COUNT 171 a 219 c 200 g 178 t
 ORIGIN
 Query Match 51.6%; Score 184.2; DB 6; Length 768;
 Best Local Similarity 80.7%; Pred. No. 4.8e-46;
 Matches 213; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTCCCGGCAATTAACAGTGGAGTCNCT 60
 |||||||
 DB 505 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCAAATACAGTGGGTCATTAT 564
 |||||||
 QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGGTCGACAAATCCACCAAGTC 120
 |||||
 DB 565 CTGCCAGAGCTGTACATGCTGAGAGAGAGAGAGTGGGCGCAACATCCACGCAAGTC 624
 |||||

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 14:11:30 ; Search time 98.6683 Seconds

(without alignments)
8148.139 Million cell updates/sec

Title: US-09-627-165E-13

Perfect score: 357

Sequence: 1 gccagattcaatccatcctt.....aaccacccgtcgatcgta 357

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N.Geneseq.101002.*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 353 | 98.9 | 357 | 22 AAC85478 | Korean mistletoe 1 |
| 2 | 230 | 64.4 | 522 | 22 AAC85479 | Korean mistletoe 1 |
| 3 | 227.4 | 63.7 | 1556 | 24 ABL56947 | Galactose recognis |
| 4 | 218 | 61.1 | 1596 | 20 AAZ09103 | Mistletoe lectin I |
| 5 | 217.6 | 61.0 | 1598 | 20 AAZ09100 | Mistletoe lectin D |
| 6 | 216.4 | 60.6 | 1923 | 18 AAT91659 | Prepro mistletoe I |
| 7 | 216.4 | 60.6 | 1923 | 20 AAV74182 | Mistletoe lectin P |
| 8 | 190.2 | 53.3 | 763 | 20 AAZ09101 | Mistletoe lectin A |
| 9 | 189 | 52.9 | 762 | 20 AAZ09104 | Mistletoe lectin A |

| | | | | | |
|----|-------|------|------|-------------|--------------------|
| 10 | 185.8 | 52.0 | 756 | 19 AAV51343 | Mistletoe rMLA var |
| 11 | 185.8 | 52.0 | 774 | 18 AAT91660 | Prepro mistletoe 1 |
| 12 | 185.8 | 52.0 | 774 | 20 AAV74180 | Mistletoe ML A-cha |
| 13 | 184.2 | 51.6 | 768 | 22 AAC85474 | A-chain gene isofo |
| 14 | 182.6 | 51.1 | 762 | 22 AAC85472 | Mistletoe rMLA DNA |
| 15 | 182.2 | 51.0 | 762 | 22 AAV51341 | A-chain gene isofo |
| 16 | 182.2 | 51.0 | 762 | 22 AAC85473 | Mistletoe lectin A |
| 17 | 179.4 | 50.3 | 768 | 20 AAZ09105 | Mistletoe lectin A |
| 18 | 90 | 25.2 | 1596 | 20 AAZ09107 | Mistletoe lectin I |
| 19 | 78.6 | 22.0 | 762 | 20 AAZ09108 | Mistletoe lectin A |
| 20 | 78.6 | 22.0 | 768 | 20 AAZ09109 | Mistletoe lectin A |
| 21 | 57 | 16.0 | 839 | 10 AAN91282 | Sequence derived f |
| 22 | 55.4 | 15.5 | 1846 | 8 AAN70526 | Sequence of cDNA 1 |
| 23 | 55.2 | 15.5 | 1825 | 22 AAT79921 | Mutant preproctin |
| 24 | 54.4 | 15.2 | 1855 | 20 AAZ04229 | PAP-256 insert con |
| 25 | 54.4 | 15.2 | 1912 | 20 AAZ04242 | PAP-284 insert DNA |
| 26 | 54.2 | 15.2 | 1837 | 22 AAT77622 | Mutant preproctin |
| 27 | 54.2 | 15.2 | 1843 | 22 AAT79937 | Mutant preproctin |
| 28 | 54.2 | 15.2 | 1849 | 22 AAT79933 | Mutant preproctin |
| 29 | 54 | 15.1 | 1855 | 20 AAZ04233 | PAP-266 insert DNA |
| 30 | 53.8 | 15.1 | 807 | 14 AAQ46085 | Sequence encoding |
| 31 | 53.8 | 15.1 | 839 | 6 AAN50381 | Estimated ricin ag |
| 32 | 53.8 | 15.1 | 1050 | 6 AAN50380 | Ricin A sequence 1 |
| 33 | 53.8 | 15.1 | 1050 | 8 AAN70519 | Ricin A gene. Esc |
| 34 | 53.8 | 15.1 | 1050 | 8 AAN70152 | Sequence of the cl |
| 35 | 53.8 | 15.1 | 1050 | 10 AAN91281 | Ricin A encoding 1 |
| 36 | 53.8 | 15.1 | 1050 | 11 AAQ05786 | Ricin A gene from |
| 37 | 53.8 | 15.1 | 1140 | 13 AAQ27876 | Partial sequence o |
| 38 | 53.8 | 15.1 | 1140 | 14 AAQ46086 | Sequence encoding |
| 39 | 53.8 | 15.1 | 1605 | 14 AAN60193 | Preproctin gene. |
| 40 | 53.8 | 15.1 | 1695 | 7 AAN60193 | Castor bean prepro |
| 41 | 53.8 | 15.1 | 1698 | 22 AAT64137 | Castor bean prepro |
| 42 | 53.8 | 15.1 | 1731 | 22 AAT64138 | Sequence of cDNA 1 |
| 43 | 53.8 | 15.1 | 1801 | 8 AAN70524 | Mutant preproctin |
| 44 | 53.8 | 15.1 | 1831 | 22 AAT79905 | Mutant preproctin |
| 45 | 53.8 | 15.1 | 1834 | 22 AAT79876 | Mutant preproctin |

ALIGNMENTS

RESULT 1
AAC85478
ID AAC85478 standard; cDNA: 357 BP.
XX
AC AAC85478:
XX
DT 16-MAY-2001 (first entry)
XX
XX Korean mistletoe lectin gene #1.
DE
XX
XX Isoform: A-chain; B-chain: biosynthesis; lectin: Korean mistletoe;
KM KML: tumour; KM-110; KML-C; KMBP; KML-ITU; KML-IL;
KW heparin binding protein; ds.
XX
XX Viscum album coloratum.
OS
XX
XX Key Location/Qualifiers
FH 55..57
FT unsure /*lag= a
FT /note= "Encodes Ser"
XX
XX EPI074560-A2.
XX
XX PD 07-FEB-2001.
XX
XX 27-JUL-2000; 2000EP-0402168.
XX
XX 27-JUL-1999; 99KR-0030638.
XX
XX (MIST-) MISTLE BIOTCH CO LTD.
XX
XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;

PI Kang T, Park C;
 XX
 DR WPI: 2001-171044/18.
 DR P-PSDB: AAB47096.

PT Novel lectin proteins isolated from Korean mistletoe, useful for
 XX enhancing immunity and effectuating anti-tumoral activity -
 PS
 XX

Claim 30; Page 33; 62pp; English.

CC The sequences given in AAC85478-79 encode lectins isolated from Korean
 CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
 CC immunity and for treating tumours. The KML's are isolated from a
 CC protein fraction derived from the leaves, stems and fruits of Korean
 CC mistletoe, which is designated KM-110. One of the isolates of Korean
 CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
 CC lymphoma.

SQ Sequence 357 BP; 77 A; 103 C; 87 G; 86 T; 4 other;

Query Match 98.9%; Score 353; DB 22; Length 357;
 Best Local Similarity 100.0%; Pred. No. 6.7e-110;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAAGTGGAGTCNTCT 60
 Db 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAAGTGGAGTCNTCT 60
 QY 61 CCACCAACAATGTACATGCTCGAGCTGGAGACGAGTTGGGGTTCGACATCCACCAAGTC 120
 Db 61 CCACCAACAATGTACATGCTCGAGCTGGAGACGAGTTGGGGTTCGACATCCACCAAGTC 120
 QY 121 CAGCAGTCCAAGATGAGCATTTTATACCAATAAGTTGAGATTTCGCCGCTAAC 180
 Db 121 CAGCAGTCCAAGATGAGCATTTTATACCAATAAGTTGAGATTTCGCCGCTAAC 180
 QY 181 TTGTGACGTTGAGCATTTTCGACGAGTGTTCAGCTGGAGTATGTTGTGAA 240
 Db 181 TTGTGACGTTGAGCATTTTCGACGAGTGTTCAGCTGGAGTATGTTGTGAA 240
 QY 241 TGCAGTGTGCGGCAATTCCTCTCCAGACCAACCTTCGCCGCTTAAGTCCGTC 300
 Db 241 TGCAGTGTGCGGCAATTCCTCTCCAGACCAACCTTCGCCGCTTAAGTCCGTC 300
 QY 301 GTGGATGCGGCCAACGATGTCACTGACATNTTCCGAACCCACCGTGCATCGTA 357
 Db 301 GTGGATGCGGCCAACGATGTCACTGACATNTTCCGAACCCACCGTGCATCGTA 357

RESULT 2
 AAC85479
 ID AAC85479 standard; cDNA: 522 BP.
 XX
 AC AAC85479;

DT 16-MAY-2001 (first entry)

XX Korean mistletoe lectin gene #2.

DE Isoform: A-chain; B-chain; Biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-III;
 KW heparin binding protein; ds.
 XX

OS Viscum album coloratum.

PN EP1074560-A2.

PD 07-FEB-2001.

PF 27-JUL-2000; 2000BP-0402168.

PR 27-JUL-1999; 99KR-0030638.

PA (MIST-) MISTLE BIOTECH CO LTD.

XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX WPI: 2001-171044/18.
 DR P-PSDB: AAB47097.

PT Novel lectin proteins isolated from Korean mistletoe, useful for
 XX enhancing immunity and effectuating anti-tumoral activity -
 PS
 XX

Claim 32; Page 34; 62pp; English.

CC The sequences given in AAC85478-79 encode lectins isolated from Korean
 CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
 CC immunity and for treating tumours. The KML's are isolated from a
 CC protein fraction derived from the leaves, stems and fruits of Korean
 CC mistletoe, which is designated KM-110. One of the isolates of Korean
 CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
 CC lymphoma.

SQ Sequence 522 BP; 112 A; 150 C; 131 G; 129 T; 0 other;

Query Match 64.4%; Score 230; DB 22; Length 522;
 Best Local Similarity 78.7%; Pred. No. 6.1e-68;
 Matches 288; Conservative 0; Mismatches 69; Indels 9; Gaps 1;

QY 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAAGTGGAGTCNTCT 60
 Db 157 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAAGTGGAGTCNTCT 216
 QY 61 CCACCAACAATGTACATGCTCGAGCTGGAGACGAGTTGGGGTTCGACATCCACCAAGTC 120
 Db 217 CTTCCGCAATGTACATGCTCGAGCTGGAGACGAGTTGGGGTTCGACATCCACCAAGTC 276
 QY 121 CAGCAGTCCAAGATGAGCATTTTATACCAATAAGTTGAGATTTCGCCGCTAAC 180
 Db 277 CAGCAGTCCAAGATGAGCATTTTATACCAATAAGTTGAGATTTCGCCGCTAAC 336
 QY 181 TTGTGACGTTGAGCATTTTCGACGAGTGTTCAGCTGGAGTATGTTGTGAA 240
 Db 337 TTGTGACGTTGAGCATTTTCGACGAGTGTTCAGCTGGAGTATGTTGTGAA 396
 QY 241 TGCAGTGTGCGGCAATTCCTCTCCAGACCAACCTTCGCCGCTTAAGTCCGTC 300
 Db 397 TGTAGGAGACCAATTCCTCTCCAGAGTGTTCGATTTGGCCGCTGATACGACCGTC 456
 QY 301 GTGGA-----TGCGCCAACAGATGTCACTGACATNTTCCGAACCCACCGTGC 351
 Db 457 TTGGAATAATAGCGGCGCTGACGAGTGTACGTGACACTCTTCGACACCGCGTGC 516
 QY 352 ATCGTA 357
 Db 517 ATCGTA 522

RESULT 3
 ABL56947

ID ABL56947 standard; DNA: 1656 BP.

AC ABL56947;

DT 08-JUL-2002 (first entry)

XX Galactose-recognising mistletoe lectin encoding polynucleotide.

XX Mistletoe; galactose-recognising mistletoe lectin; MLIII; gene; ds.

OS Viscum album.

FT Key Location/Qualifiers

FT CDS 1..1656
 FT /*tag= a

QY 301 GTGATGCGGCAACGATGTCACCTGACACTTTTCCGAACCCAGCGTGGCATGCT 356
 Db 799 AT-----AGCCGATGATGTTACTGTCAGTGTCTTCCGAACTACAGTGGCGATTGT 848

RESULT 5

AAZ09100
 ID AAZ09100 standard; DNA; 1598 BP.

XX AAZ09100;

XX 18-OCT-1999 (first entry)

XX Mistletoe lectin DNA consensus sequence.

XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; ss.

XX Viscum album.

XX DE19804210-A1.

XX 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.

XX 03-FEB-1998; 98DE-1004210.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Morris P, Stiefel T, Voelter W, Walters P;

XX WPI; 1999-445335/38.

XX Preparation of mistletoe lectins in heterologous systems.

XX Particularly for use as anticancer agents and immunostimulants

XX Claim 11; Page 32-33; 78pp; German.

XX This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and (I) and its
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of mistletoe lectin DNA described in the
 CC specification.

XX Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;

Query Match

XX Best Local Similarity 61.0%; Score 217.6; DB 20; Length 1598;

XX Matches 266; Conservative 13; Mismatches 71; Indels 6; Gaps 1;

QY 1 GCCAGATTCATCCCATNTGTGAGGCTTCCGCGCAATTAACAGTGGGAGTCNTCT 60

Db 500 GCCAGATTCATCCCATNTGTGAGGCKMYCGCAARAVYTAACAGTGGGAGTCNTCT 60

QY 61 CCACCAACATGATGATGCTGAGAGCGAGTGGGGTGACATCCACCAAGTC 120

Db 560 CTGGCAGACGCTGTACATGCTGAGAGCGAGTGGGGCCAAACATCCACCAAGTC 619

QY 121 CAGCAGTCCAGAGATGTCATTTTAAATACCAATAGATTTGAGATTTCCGGCGTAC 180

Db 620 CAGCATTCAACGATGGCGTTTAAATACCAACATTCGTTGGCTATATTCMCTGCTAAC 679

QY 181 TTGTGAGGNTGAGCATGTTGCGAGTGATGTCAGCTTGGGATCATGTTGCGAA 240
 Db 680 TTGTGAGGNTGAGCATGTTGCGAGTGATGTCAGCTTGGGATCATGTTGCGAA 240
 QY 241 TGCAGTGGTGGCCATTTCTCTCTCTGACACACCTTGGCGGTCCTTAAGTCCGTC 300
 Db 740 TCCGAGAGCGGCCATCTCTCTGACGTCGCTATTGGCCGCTGGTCAACGACCGTG 799
 QY 301 GTGATGCGGCAACGATGTCACCTGACACTTTTCCGAACCCAGCGTGGCATGCT 356
 Db 800 AT-----AGCCGATGATGTTACTGTCAGTGTCTTCCGAACTACAGTGGCGATTGT 849

RESULT 6

AAAT91659
 ID AAAT91659 standard; DNA; 1923 BP.

XX AAAT91659;

XX 18-DEC-1997 (first entry)

XX Prepro mistletoe lectin.

XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss.

XX Viscum album.

XX Key Location/Qualifiers

XX CDS 55..1749

XX misc_feature 1..204

XX misc_feature /tag= b

XX misc_feature /note= "ML gene fragment h"

XX misc_feature /tag= c

XX misc_feature /note= "ML gene fragment f"

XX misc_feature /tag= d

XX misc_feature /note= "ML gene fragment e"

XX misc_feature /tag= e

XX misc_feature /note= "ML gene fragment g"

XX misc_feature /tag= f

XX misc_feature /note= "ML gene fragment i"

XX EP751221-A1.

XX 02-JAN-1997.

XX 26-JUN-1995; 95EP-0109949.

XX 26-JUN-1995; 95EP-0109949.

XX (MADU) MADDAUS KOEHN AG.

XX Baur A, Eck J, Lentzen H, Zinke H;

XX WPI; 1997-054678/06.

XX P-PSDB; 10021.

XX Nucleic acid encoding pre-pro form of mistletoe lectin - for

XX therapeutic or diagnostic use

XX Claim 1; Fig 4c; 30pp; German.

XX Mistletoe lectin is a cytotoxic agent that has been used for tumour
 CC therapy. It can be used in immunotoxins and medicaments. Nucleic
 CC acid fragments can be used in diagnostic methods. Mistletoe lectin
 CC (AAAT91659) comprises an A chain (AAAT91660) and a B chain (AAAT91661).
 XX Sequence 1923 BP; 465 A; 488 C; 495 G; 475 T; 0 other;

| | | | | |
|-----------------------|--------------|--------------------|----------------|--------------|
| Query Match | 60.68; | Score 216.4; | DB 18; | Length 1923; |
| Best Local Similarity | 77.28; | Pred. No. 4.7e+63; | | |
| Matches 275; | Conservative | 0; | Mismatches 75; | Indels 6; |
| | | | | Gaps 1 |

| | | | |
|----|-----|---|------|
| Qy | 1 | GGCATTCAATCCCATCTGTTGGAGGCTTCCGCGGCAATTAACAGTGGGGAGTCTCT | 60 |
| | | | |
| | | | |
| | | | |
| Db | 652 | GGCATTCAATCCCATCTTATGGAGGGCTCCCAATACATTAAACAGTGGGGCTCATTT | 711 |
| | | | |
| | | | |
| | | | |
| Qy | 61 | CCACAAACATGTTCATGTCTCGAGCTGGAGAGAGTGGGGTGCACATCCACCAAGTC | 120 |
| | | | |
| | | | |
| | | | |
| Db | 712 | CTCCAGAGGTGTACATCTGGAGCTGGAGACCAATTTGGGGCCAACTCCACCAAGTC | 771 |
| | | | |
| | | | |
| | | | |
| Qy | 121 | CACAGCTCCACAGATGGCATTTTAAATACCAATAAAGATTGCAGATTTCCGCGGTAC | 180 |
| | | | |
| | | | |
| | | | |
| Db | 772 | CACATTTCAACGATGGGGTTTTTAATAACCAATTCGGTTGGCTATACCCCCCGGTAC | 831 |
| | | | |
| | | | |
| | | | |
| Qy | 181 | TTTGTGACGNTGAGCAATGTTCCGCGAGCTATCTCCAGCTTGGCANTCATGTTGTGAA | 240 |
| | | | |
| | | | |
| | | | |
| Db | 832 | TTGTGTGACGTTGACCAATGTTCCGCGAGCTGATGCGCAGTTGGCATCATGTTGTTGTA | 891 |
| | | | |
| | | | |
| | | | |
| Qy | 241 | TGCAGTGTGCGGCATTTCTCTCTCTGACACACCTTCGCGCGTGCCTTAAGTCCGTC | 300 |
| | | | |
| | | | |
| | | | |
| Db | 892 | TGCGGAGAGCGGCATCTTCTCTGAGGTGCGCTATTGGCCGTGGTCAATAGAACCGTG | 951 |
| | | | |
| | | | |
| | | | |
| Qy | 301 | GTGATGTCGGCCAGACGTGTCACCTGCACACTTTTCCAAACCCACGCTGGCATCTGT | 356 |
| | | | |
| | | | |
| | | | |
| Db | 952 | AT-----ACCGCATGATTTACCTGCGAGTCTTCCGAACCTACGCTGGCGGATTTGT | 1001 |

| RESULT 7 |
|------------------------------------|
| AAV74182 |
| ID AAV74182 standard; DNA; 1923 BP |

| | | |
|----|---|-----------------------------|
| DT | 30-APR-1999 | (first entry) |
| XX | | |
| DE | Mistletoe lectin prepro-protein DNA. | |
| XX | | |
| KW | ML; mistletoe; lectin; ML; transgenic plant; glycosylation; | |
| KW | dimer; immunotoxin; large-scale production; diagnosis; therapeutic; | |
| KW | cancer; ss. | |
| XX | | |
| OS | viscum album. | |
| XX | | |
| PH | key | location/Qualifiers |
| FT | CDS | 55..1749 |
| FT | | /tag="a |
| FT | | /product="mistletoe lectin" |

| | | |
|----|---------------------------|---------------|
| PN | EP8884388-A1. | |
| XX | 16-DEC-1998. | |
| XX | | |
| XX | 26-JUN-1995; | 95EP-0109949. |
| XX | | |
| PR | 26-JUN-1995; | 95EP-0109949. |
| PR | 26-JUN-1995; | 98EP-0105660. |
| PA | | |
| XX | (MADU) MADDAUS KOELN AG. | |

aa Baur A, Eck J, Lentzen H, Zinke H
pi
xx WPI: 1999-026582/03.
DR
DR P-PSDB; AAW90127.

XX New transgenic plant expressing mistletoe lectin - useful for
PT producing recombinant lectin in e.g. cancer diagnosis and therapy
XX
PS Claim 1a; Fig 4c; 30pp; German.

vector capable of encoding a mistletoe (Viscum album) lectin
CC preprotein or a biologically active fragment. The specification
CC also describes a polypeptide produced by a plant where the polypeptide
CC exhibits at least one enzymatic modification other than the glycosylation
CC that occurs in Viscum album or the polypeptide is a fusion protein, a
CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
CC polypeptide or the polypeptide dimer. The plants are used for large-scale
CC production of mistletoe lectin for diagnostic or therapeutic purposes
CC (e.g. in cancer therapy). This sequence encodes the mistletoe lectin
CC used in the method of the invention.

XX Sequence 1923 BP; 465 A; 483 C; 495 G; 475 T; 0 other;
SQ

| | | | | |
|-----------------------|--------------|--------------------|----------------|--------------|
| Query Match | 60.6%; | Score 216.4; | DB 20; | Length 1923; |
| Best Local Similarity | 77.2%; | Pred. No. 4.7e+63; | | |
| Matches 275; | Conservative | 0; | Mismatches 75; | Indels 6; |
| | | | | Gaps 1; |

| | | | |
|----|-----|--|------|
| Oy | 1 | GCCAGATTCAATCCCATCTGTGGAGCCTTGCCCGCAAAATTAAAGTGGGGGACCTTC | 60 |
| Db | 652 | GCCAGATTCAATCCCATCTTAATGAGGGCTGCCAATAATCAATTAAAGTGGGGCTATT | 711 |
| Oy | 61 | CCACCACAATATATATCTGAGGTGGAGCGAGTTGGGGTGCACATCACCCAAGTC | 120 |
| Db | 712 | CCTCCAGAGCTTAATATCTGGAGGTGAGACGAATTGGGCCAACAAATCACGAGAATC | 771 |
| Oy | 121 | CAGCAGTCCAGATGGCATTTTAAATACCCAAATTAAGATTGCAGATTTCCGCCGGTAAC | 180 |
| Db | 772 | CACCATTCACCGATGGCGTTTTTAATAAACCAAATTCGTTAGCTATACCCCCCGGTAAAC | 831 |
| Oy | 181 | TTTTGTAGCAGNTAGCAATGTTGCCACGATCTCCACCTTGGGGATCAATGTTTGGAA | 240 |
| Db | 832 | TTCGTGACGTTGACCAATGTTGCACAGCGAATGCCACCTTGGGATCATATGTTTGGTA | 891 |
| Oy | 241 | TGCAGTGTGTGGCATTTCTCTCTCGACACACCCTTGGCCGCTGCTCTTAAGTCCGTC | 300 |
| Db | 892 | TGCGGAGAGGGGCCATCTTCCTCTGAGGTGCCCTATTGGCCGCTGTATACGAACCGTG | 951 |
| Oy | 301 | GTCGATCGGCCCAACGATGTCACCTGATCTTTTCCGAACCCACCGTGCAGATCGT | 356 |
| Db | 952 | AT-----ACCCGATGATGTACTCTCAATGCTGTGGAACTACAGGTGCGGATGTGT | 1001 |

| | |
|----------|--------------------------------|
| RESULT | 8 |
| AAZ09101 | |
| ID | AAZ09101 standard; DNA; 763 BP |

18-OCT-1999 (first entry)

XX Morris P, Stiefel T, Voelter W, Welters P,
PI
XX
DR WPI; 1999-445335/38.

CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumor and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrollably cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
CC described in the specification.

| | | | | |
|-----------------------|------------------|--------------------|--------|-------------|
| Query Match | 53.38; | Score 190.2; | DB 20; | Length 763; |
| Best Local Similarity | 78.48; | Pred. No. 2.6e-54; | | |
| Matches 207; | Conservative 13; | Mismatches 44; | T-2.37 | 2 |

| | |
|----------|---------------------------------|
| RESULT 9 | |
| AAZ09104 | |
| ID | AAZ09104 standard; DNA; 762 BP. |
| XY | |

Mistletoe lectin A1 DNA fragment.

ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response; lymphokine-producing macrophage; uncontrolled cell growth; treatment; cancer; cytotoxicity; antigen; isoform; lectin A1; ds.

Viscum album.

DE19804210-A1.

12-AUG-1999

03-FEB-1998; 98DE-1004210.

03-FEB-1998; 98DE-1004210.

(BIOS-) BIOSYN ARZNEIMITTEL GMBH.

DR WPI; 1999-445335/38
DR P-PSDB; AAY25980.

preparation of mistletoe lectins in heterologous systems,
particularly for use as anticancer agents and immunostimulants

rs cladii 15; flg 2A; 78pp; German
XX

CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity (II) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC encodes a fragment of the mistletoe lectin A1 protein.
XX

Sequence 762 BP; 171 A; 230 C; 183 G; 178 T; 0 other;

| | | | | |
|-----------------------|-----------------|--------------------|--------|-------------|
| Query Match | 52.98; | Score 189; | DB 20; | Length 762; |
| Best Local Similarity | 81.88; | Pred. No. 6.6e-54; | | |
| Matches 216; | Conservative 0; | Mismatches 48; | | |

Accuracy 01.00; Pred. NO. 6.6e-54;
Matches 216; Conservative 0; Mismatches 48. Indels 0 / 0

[illegible]

RESULT 10
AAV51343

| ID | standard; DNA; 756 BP |
|----------|-----------------------|
| AAV51343 | |
| XX | |

AC AAV51343
xy

| | | |
|----|-------------|---------------|
| DT | 23-OCT-1998 | (first entry) |
| XY | | |

DE Mistletoe rMLA variant DNA.
XX

AM Lecithin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
KW Intracellular; processes; metabolism

autoimmune disease: allergy; tumour; cell proliferation; activation; immunisation; treatment; disorder

OS *Viscum album.*

| FH | Key | Location/Qualifiers |
|----|-----|---------------------|
| | | |

```
/*tag= a
```

```

/product= "lectin A-chain"

```

FT /note="Partial sequence"

XX MO9829540-A2.

PN 09-JUL-1998.

XX 02-JAN-1998; 98WO-EP00009.

XX 02-JAN-1997; 97EP-0100012.

XX (BRAT-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.

XX Eck J, Schmidt A, Zinke H;

XX WPI: 1998-388122/33.

XX Nucleic acid encoding fusion protein containing mistletoe lectin A

PT chain - useful for treatment of proliferative and autoimmune

PT diseases, allergies and tumours

XX Disclosure; Fig 11a'; 115pp; German.

XX This sequence encodes a variant mistletoe lectin A-chain, rMLA. This

CC sequence can be used in the construction of a fusion protein which

CC comprises an effector module that is cytotoxic intracellularly, a

CC processing module covalently bonded to the effector module and

CC containing a protease recognition sequence, and a targeting module

CC covalently bonded to the processing module, able to bind specifically to

CC the surface of a cell so as to mediate internalisation of the fusion

CC protein. Such a fusion protein can be used for treating disorders

CC involving proliferation and/or elevated activation of cells, especially

CC autoimmune disease, allergy and tumours. The proteins can be administered

CC e.g. by injection or topically but especially by intravenous injection,

CC at 1 ng to 500 mg/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.

CC Fusion proteins can develop toxic activity in a wide range of target

CC cells. The processing module prevents extracellular dissociation, and

CC fusion proteins based on mistletoe lectin A-chain are far more active

CC than those based on ricin and do have the associated problems of

CC non-specific toxicity. The protein may be expressed in a non-glycosylated

CC form that does not bind to sugar receptors in the liver, and which has a

CC long half-life in the blood. Where the mistletoe lectin B-chain is used,

CC it actively assists in translocation of the ML A-chain from the

CC endoplasmic reticulum to the cytoplasm.

CC

XX Sequence 756 BP; 167 A; 229 C; 181 G; 179 T; 0 other;

XX

Query Match 52.0%; Score 185.8; DB 19; Length 756;

Best Local Similarity 82.4%; Pred. No. 8.1e-53;

Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

XX 1 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGGCAATTAACAGTGGAGTCTCT 60

DB 499 GCCAGATTCAATCCCATCTTATGAGGCTCCCAATTAACAGTGGAGGCTCATTT 558

XX 61 CCACAAACATGTACATGCTCGAGCTGAGAGAGTGGGGTGCACATCCACCAAGTC 120

DB 559 CTGCCAGACTGTACATGCTGGAGCTGGAGAGAGTGGGGCAACAATCCACCAAGTC 618

XX 121 CAGCAGTCCAAAGATGGCATTTTATATACCAATTAAGTTCGACATTTCCCGCGTAC 180

DB 619 CAGATTCAACCAAGTGGCGTTTATTAACCAATTCGATGGTCTACATCCCGCGTAC 678

XX 181 TTTGTGACGNTGAGCAATGTTCCGACGTGATCTCCAGCTTGGCGATCATGTTTGC 240

DB 679 TTGCTGACGTTGACCAATGTTCCGACGTGATCTCCAGCTTGGCGATCATGTTTGA 738

XX 241 TGCAGTGTGGCCAT 256

DB 739 TGCAGAGAGCGGCCAT 754

RESULT 11

AAT91660

ID AAT91660 standard; DNA: 774 BP.

XX AAT91660;

AC 18-DEC-1997 (first entry)

XX Prepro mistletoe lectin A chain.

XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss.

XX Viscum album.

XX Key Location/Qualifiers

FT CDS 4..765

FT /*tag= a

XX EP751221-A1.

XX 02-JAN-1997.

XX 26-JUN-1995; 95EP-0109949.

XX 26-JUN-1995; 95EP-0109949.

XX (MADU) MADUUS KOELN AG.

XX Baur A, Eck J, Lentzen H, Zinke H;

XX WPI: 1997-054678/06.

XX P-PSDB; AAM10022.

XX Nucleic acid encoding pre-pro form of mistletoe lectin - for

PT therapeutic or diagnostic use

XX Claim 2; Fig 4A; 30pp; German.

XX Mistletoe lectin is a cytotoxic agent that has been used for tumour

CC therapy. It can be used in immunotoxins and medicaments. Nucleic

CC acid fragments can be used in diagnostic methods. Mistletoe lectin (

CC AAT91659) comprises an A chain (AAT91660) and a B chain (AAT91661).

XX

XX Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other;

XX

Query Match 52.0%; Score 185.8; DB 18; Length 774;

Best Local Similarity 82.4%; Pred. No. 8.2e-53;

Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

XX 1 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGGCAATTAACAGTGGAGTCTCT 60

DB 505 GCCAGATTCAATCCCATCTTATGAGGCTCCCAATTAACAGTGGAGGCTCATTT 564

XX 61 CCACAAACATGTACATGCTCGAGCTGAGAGAGTGGGGTCCACATTCACCAAGTC 120

DB 565 CTGCCAGACTGTACATGCTGGAGCTGGAGAGAGTGGGGCAACAATTCACCAAGTC 624

XX 121 CAGCAGTCCAAAGATGGCATTTTATATACCAATTAAGTTCGACATTTCCCGCGTAC 180

DB 625 CAGCATTCAACCAAGTGGCGTTTATTAACCAATTCGATGGTCTACATCCCGCGTAC 684

XX 181 TTTGTGACGNTGAGCAATGTTCCGACGTGATCTCCAGCTTGGCGATCATGTTTGC 240

DB 685 TTGCTGACGTTGACCAATGTTCCGACGTGATCTCCAGCTTGGCGATCATGTTTGA 744

XX 241 TGCAGTGTGGCCAT 256

DB 745 TGCAGAGAGCGGCCAT 760

RESULT 12

AAV74180

ID AAV74180 standard; DNA: 774 BP.

XX AAV74180;

XX 30-APR-1999 (first entry)
 DT Mistletoe ML A-chain DNM.
 XX
 DE
 XX
 KW ML: mistletoe; lectin; MLA; A-chain; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer; ss.
 XX
 OS Viscum album.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 4..765
 FT /*tag= a
 FT /product= "mistletoe lectin"
 XX
 PN EP884388-A1.
 XX
 PD 16-DEC-1998.
 XX
 PE 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 98EP-0105660.
 XX
 PA (MADU) MADUS KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX
 DR WPI; 1999-026582/03.
 DR P-PSDB: AAM90125.
 XX
 PT New transgenic plant expressing mistletoe lectin - useful for
 PT producing recombinant lectin in e.g. cancer diagnosis and therapy
 XX
 PS Claim 2; Fig 4a; 30pp; German.
 XX
 CC This invention describes a novel transgenic plant transformed with a
 CC vector capable of encoding a mistletoe (Viscum album) lectin
 CC preproprotein or a biologically active fragment. The specification
 CC also describes a polypeptide produced by a plant where the polypeptide
 CC exhibits at least one enzymatic modification other than the glycosylation
 CC that occurs in Viscum album or the polypeptide is a fusion protein, a
 CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 CC polypeptide or the polypeptide dimer. The plants are used for large-scale
 CC production of mistletoe lectin for diagnostic or therapeutic purposes
 CC (e.g. in cancer therapy). This sequence encodes the mistletoe lectin
 CC A-chain which is contained in expression vector pT7MA.
 CC
 SQ Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other:
 Query Match 52.0%; Score 185.8; DB 20; Length 774;
 Best Local Similarity 82.4%; Pred. NO. 8.2e-53;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAATTCATCCCATNTGTGAGGCTTCGCCGAATTAAGTGGGGAGTCNTCT 60
 Db 505 GCCAATTCATCCCATNTGTGAGGCTTCGCCGAATTAAGTGGGGAGTCNTCT 60
 QY 61 CCACCAACATGTACATCTGAGCTGAGAGCGATGGGGGTGACAAATCCACCAAGTC 120
 Db 565 CTGCGACAGCTGTACATCTGAGCTGAGAGCGATGGGGGTGACAAATCCACCAAGTC 120
 QY 121 CAGCAGTCCAAAGATGGCATTTTAATACCAATAAGATTGAGATTTCCCGCGTAAC 180
 Db 625 CAGCAGTCCAAAGATGGCATTTTAATACCAATAAGATTGAGATTTCCCGCGTAAC 180
 QY 181 TTGTGACGNTGAGCAATGTTTCGAGAGTATCCAGCTTGGCGATCATGTTGTGAA 240
 Db 685 TTGTGACGNTGAGCAATGTTTCGAGAGTATCCAGCTTGGCGATCATGTTGTGAA 240
 QY 241 TGCAGTGTGGGCGCAT 256
 Db 744 TGCAGTGTGGGCGCAT 256

Db 745 TGCAGGAGACGGCCCAT 760
 RESULT 13
 AAC85474
 ID AAC85474 standard; cDNA; 768 BP.
 XX
 AC AAC85474;
 XX
 AC
 XX
 DT 16-MAY-2001 (first entry)
 DE
 XX
 DE A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #3.
 XX
 KW Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-III;
 KW heparin binding protein; ds.
 XX
 OS Viscum album coloratum.
 XX
 PN EP1074560-A2.
 XX
 PD 07-FEB-2001.
 XX
 PE 27-JUL-2000; 2000EP-0402168.
 XX
 PR 27-JUL-1999; 99KR-0030638.
 XX
 PA (MIST-) MISTLE BIOTECH CO LTD.
 XX
 PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX
 DR WPI; 2001-171044/18.
 DR P-PSDB: AAB47092.
 XX
 PT Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 XX
 PS Claim 5; Page 27-28; 62pp; English.
 XX
 CC The sequences given in AAC85472-74 encode isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumors. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-M125 lymphoma.
 CC
 SQ Sequence 768 BP; 171 A; 219 C; 200 G; 178 T; 0 other:
 Query Match 51.6%; Score 184.2; DB 22; Length 768;
 Best Local Similarity 80.7%; Pred. NO. 2.9e-52;
 Matches 213; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1 GCCAATTCATCCCATNTGTGAGGCTTCGCCGAATTAAGTGGGGAGTCNTCT 60
 Db 505 GCCAATTCATCCCATNTGTGAGGCTTCGCCGAATTAAGTGGGGAGTCNTCT 60
 QY 61 CCACCAACATGTACATCTGAGCTGAGAGCGATGGGGGTGACAAATCCACCAAGTC 120
 Db 565 CTGCGACAGCTGTACATCTGAGCTGAGAGCGATGGGGGTGACAAATCCACCAAGTC 120
 QY 121 CAGCAGTCCAAAGATGGCATTTTAATACCAATAAGATTGAGATTTCCCGCGTAAC 180
 Db 625 CAGCAGTCCAAAGATGGCATTTTAATACCAATAAGATTGAGATTTCCCGCGTAAC 180
 QY 181 TTGTGACGNTGAGCAATGTTTCGAGAGTATCCAGCTTGGCGATCATGTTGTGAA 240
 Db 685 TTGTGACGNTGAGCAATGTTTCGAGAGTATCCAGCTTGGCGATCATGTTGTGAA 240
 QY 241 TGCAGTGTGGGCGCATTCCTCT 264
 Db 745 TGCAGGAGACGGCATTCCTCT 768


```
Db 952 AT-----AGCCGATGATGTACTGACGTGCTTGGAAACCTACGCTGGCATTTGT 1001
;
; Sequence 38, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776,059B
; EARLIER FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Viscum album
US-08-776-059-38

Query Match 52.0%; Score 185.8; DB 4; Length 705;
Best Local Similarity 82.4%; Pred. No. 1e-55;
Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCMTGTGAGGCTTCGCCGCAATTAACAGTGGAGTCNTCT 60
Db 448 GCCAGATTCAATCCCATCTATGTAGAGGCTTCGCCAATTAACAGTGGAGTCNTCT 507
QY 61 CCACCAACATGTACATGCTGAGCGTGGAGACAGATTTGGGCTGACAAATCCACCAAGTC 120
Db 508 CTGCGACAGCTGTACATGCTGAGCGTGGAGACAGATTTGGGCGCAACATCCACGCAAGTC 567
QY 121 CACGAGTCCAAAGATGGCATTTTAAATACCAATTAAGATTGAGATTTCCGCGGTAAAC 180
Db 568 CACGATTCACACGATGGCGTTTAAATACCAATTCGTTGCTATACCCCGGTAAAC 627
QY 181 TTGTGACGNTGAGCAATGTTCCGACGATGATCTCCAGCTTGCGCATGTTGTTGCA 240
Db 628 TTGTGACGNTGAGCAATGTTCCGACGATGATCTCCAGCTTGCGCATGTTGTTGTA 687
QY 241 TGCAGTGTGCGCCAT 256
Db 688 TGCAGTGTGCGCCAT 703

RESULT 3
US-08-776-059-30
; Sequence 30, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776,059B
; EARLIER FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 30
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Viscum album
US-08-776-059-30

Query Match 52.0%; Score 185.8; DB 4; Length 774;
Best Local Similarity 82.4%; Pred. No. 1e-55;
Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCMTGTGAGGCTTCGCCGCAATTAACAGTGGAGTCNTCT 60
Db 505 GCCAGATTCAATCCCATCTATGTAGAGGCTTCGCCAATTAACAGTGGAGTCNTCT 564
QY 61 CCACCAACATGTACATGCTGAGCGTGGAGACAGATTTGGGCTGACAAATCCACCAAGTC 120
Db 565 CTGCGACAGCTGTACATGCTGAGCGTGGAGACAGATTTGGGCGCAACATCCACGCAAGTC 624
QY 121 CACGAGTCCAAAGATGGCATTTTAAATACCAATTAAGATTGAGATTTCCGCGGTAAAC 180
Db 625 CACGATTCACACGATGGCGTTTAAATACCAATTCGTTGCTATACCCCGGTAAAC 684
QY 181 TTGTGACGNTGAGCAATGTTCCGACGATGATCTCCAGCTTGCGCATGTTGTTGCA 240
Db 685 TTGTGACGNTGAGCAATGTTCCGACGATGATCTCCAGCTTGCGCATGTTGTTGTA 744
QY 241 TGCAGTGTGCGCCAT 256
Db 745 TGCAGTGTGCGCCAT 760

RESULT 4
US-08-356-786-7
; Sequence 7, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
```


STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..807
OTHER INFORMATION: /note="product = "Ricin-A chain"
US-08-356-786-7

Query Match 15.1%; Score 53.8; DB 2; Length 807;
Best Local Similarity 50.0%; Pred. No. 3.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTGTCGAGCGCTTCCGCGCAAAATTAACAGTGGGAGTCTCT 60
DB 538 GCAGATTCAATATATTTGAGGAGAAATGCGCAGAGATTAAGTACACCGAGATCT 597
QY 61 CCACCAACATGTACATGCTGAGCTGGAGACGAGAGTGGGTCGACAAATCCACCAAGTC 120
DB 598 GCACGAGATCTAGCGTAACTTACACTTGAATAGTGGGAGACTTTCACACTGCATTT 657
QY 121 CAGCACTCCAGAGATGGCATTTTAAATCCCAATAGATTGACAGATTTCCGCGGTAC 180
DB 658 CAAGAGCTTACCAAGAGACCTTTGCTAGTCCAAATTCACACTGCAAAAGAGTAATGTTCC 717
QY 181 TTGTGACGNTGAGCAATGTTGCGACGATGATCTCCAGCTTGGCGATCATGTTGTTGAA 240
DB 718 AATTCAGTGTACAGTATGAGTATATTAATCCCTATCATAGCTCATGTTGTTATAGA 777
QY 241 TGCAGTGTGCGCCATTCTC 260
DB 778 TGGCAGACTCCACCATCTGC 797

RESULT 5

US-08-218-303-15
Sequence 15, Application US/08218303
Patent No. 5547867
GENERAL INFORMATION:
APPLICANT: Kara, Bhuphendra V.
APPLICANT: Hockney, Robert C.
APPLICANT: Fliton, John E.
TITLE OF INVENTION: FERMENTATION PROCESS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,303
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/841,533
FILING DATE: 26-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/3893/94908/JMW.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 143..943
US-08-218-303-15

Query Match 15.1%; Score 53.8; DB 1; Length 1140;
Best Local Similarity 50.0%; Pred. No. 4.2e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTGTCGAGCGCTTCCGCGCAAAATTAACAGTGGGAGTCTCT 60
DB 677 GCAGATTCAATATATTTGAGGAGAAATGCGCAGAGATTAAGTACACCGAGATCT 736
QY 61 CCACCAACATGTACATGCTGAGCTGGAGACGAGAGTGGGTCGACAAATCCACCAAGTC 120
DB 737 GCACGAGATCTAGCGTAACTTACACTTGAATAGTGGGAGACTTTCACACTGCATTT 796
QY 121 CAGCACTCCAGAGATGGCATTTTAAATCCCAATAGATTGACAGATTTCCGCGGTAC 180
DB 797 CAAGAGCTTACCAAGAGACCTTTGCTAGTCCAAATTCACACTGCAAAAGAGTAATGTTCC 856
QY 181 TTGTGACGNTGAGCAATGTTGCGACGATGATCTCCAGCTTGGCGATCATGTTGTTGAA 240
DB 857 AATTCAGTGTACAGTATGAGTATATTAATCCCTATCATAGCTCATGTTGTTATAGA 916
QY 241 TGCAGTGTGCGCCATTCTC 260
DB 917 TGGCAGACTCCACCATCTGC 936

RESULT 6

US-08-338-793D-60
Sequence 60, Application US/08338793D
Patent No. 5840521
GENERAL INFORMATION:
APPLICANT: Barth, Peter Thomas
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY CUSHMAN
ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
ADDRESSEE: PULSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,793D
FILING DATE: 08-NO. 5840521-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842,081
FILING DATE: 26-FEB-92
CLASSIFICATION: 435
APPLICATION NUMBER: 9104017.0
FILING DATE: 26-FEB-91
APPLICATION NUMBER: 9109188.4
FILING DATE: 29-APR-91
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-338-793D-60

Query Match 15.1%; Score 53.8; DB 2; Length 1140;
Best Local Similarity 50.0%; Pred. No. 4.2e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATGCTGAGAGGCTTCCGCGCAATTAACAGTGGAGTCNTCT 60
Db 677 GCAAGATTCATATATTTAGGAGGAGAAATGCGCAGAGATTAGGTACAAACCGAGATCT 736
QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGGTGACAAATCCACCAAGTC 120
Db 737 GCACCGATTCCTAGCTTAATTAACCTGAGATAGTGGGAGACTTCCACTGCAAT 796
QY 121 CACGAGTCCAGAGATGGCAATTTTAAATACCAATTAAGATTGACAGATTCCGCGGTAA 180
Db 797 CAAGATCTACCAAGAGACCTTTCTAGTCCAAATTAACCAAGAGATGATGTTCC 856
QY 181 TTTGTAGCAGTGAAGCAATGTTGCGAGCTGATCTCCAGCTTGGCGATGTTGTGGA 240
Db 857 AAATTCAGTGTAGCATGTGATTAATTAATCCATATCAATGCTCTCATGTATAGA 916
QY 241 TGCAGTGTGCGGCATTCCTC 260
Db 917 TCGCAGCTTCACCAATCGTC 936

RESULT 7

US-08-356-786-9
Sequence 9, Application US/08356786
Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1605
OTHER INFORMATION: /note= "product = "G-Fir""

Query Match 15.1%; Score 53.8; DB 2; Length 1605;
Best Local Similarity 50.0%; Pred. No. 5.1e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATGCTGAGAGGCTTCCGCGCAATTAACAGTGGAGTCNTCT 60
Db 544 GCAAGATTCATATATTTAGGAGGAGAAATGCGCAGAGATTAGGTACAAACCGAGATCT 603
QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGTGGGTGACAAATCCACCAAGTC 120
Db 604 GCACCGATTCCTAGCTTAATTAACCTGAGATAGTGGGAGACTTCCACTGCAAT 663
QY 121 CACGAGTCCAGAGATGGCAATTTTAAATACCAATTAAGATTGACAGATTCCGCGGTAA 180
Db 664 CAAGATCTACCAAGAGACCTTTCTAGTCCAAATTAACCAAGAGATGATGTTCC 723
QY 181 TTTGTAGCAGTGAAGCAATGTTGCGAGCTGATCTCCAGCTTGGCGATGTTGTGGA 240
Db 724 AAATTCAGTGTAGCATGTGATTAATTAATCCATATCAATGCTCTCATGTATAGA 783
QY 241 TGCAGTGTGCGGCATTCCTC 260
Db 784 TCGCAGCTTCACCAATCGTC 803

RESULT 8

US-09-147-208-33
Sequence 33, Application US/09147208
Patent No. 6333303

GENERAL INFORMATION:

APPLICANT: Antiviral Ricin-Like Proteins
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/147,208
FILING DATE: 02-MAR-1999
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Rudolph, John R.
REGISTRATION NUMBER: 38,003
REFERENCE/DOCKET NUMBER: 7841-76
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:

61 CCACCAACATGTACATGCTCGAGCTGAGAGACGAGTTGGGGTGCACAATCCACCAAGTC 120

Db 706 GCACCGAGTCTAGCCTAATTACCTTGAGATAGTGGGAGACTTCCACTCAATT 765
QY 121 CAGCAGTCCAAAGATGATTTTAAATACCAAAATTAAGTTGCAATTTCCGGGTAC 180
Db 766 CAGAGTCTAACCAGAGACCTTTGCTAGTCCAAATTCAGTCAAGAGTAAATGTTCC 825
QY 181 TTGTGACGNTGAGCAATGTTGGCAGCGATCTCCAGCTTGGCGATGTTGTGCA 240
Db 826 AAATTCAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 885
QY 241 TGCAGTGTGGCCATCTC 260
Db 886 TGGCAGCTCCACCATCTC 905

RESULT 11
US-09-147-208-54
; Sequence 54, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1855 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-54

Query Match 15.1%; Score 53.8; DB 4; Length 1855;
Best Local Similarity 50.0%; Pred. No. 5.4e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 1 GCCAATTCATCCCATGNTGTGGAGGCTTCCGCGCAATTAACAGTGGGAGTCNTCT 60
Db 646 GCACATTCCTATATATGAGGAGAAATGCGCAGAGAAATTAAGTCAACCGAGATCT 705
QY 61 CCACAAACATGTACATCTCGAGCTGAGAGAGAGTTGGGGTGCACAAATCCACCAAGTC 120
Db 706 GCACAGATCTAGGTATTAACACTTGAGATATGTTGGGGAGACTTTCACCTGCAATT 765
QY 121 CAGCAGTCCAAAGATGATTTTAAATACCAAAATTAAGTTGCAATTTCCGGGTAC 180
Db 766 CAGAGTCTAACCAGAGACCTTTGCTAGTCCAAATTCAGTCAAGAGTAAATGTTCC 825
QY 181 TTGTGACGNTGAGCAATGTTGGCAGCGATCTCCAGCTTGGCGATGTTGTGCA 240

Db 826 AAATTCAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 885
QY 241 TGCAGTGTGGCCATCTC 260
Db 886 TGGCAGCTCCACCATCTC 905

RESULT 12
US-09-147-208-23
; Sequence 23, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-23

Query Match 15.1%; Score 53.8; DB 4; Length 1879;
Best Local Similarity 50.0%; Pred. No. 5.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 1 GCCAATTCATCCCATGNTGTGGAGGCTTCCGCGCAATTAACAGTGGGAGTCNTCT 60
Db 670 GCACATTCCTATATATGAGGAGAAATGCGCAGAGAAATTAAGTCAACCGAGATCT 729
QY 61 CCACAAACATGTACATCTCGAGCTGAGAGAGAGTTGGGGTGCACAAATCCACCAAGTC 120
Db 730 GCACAGATCTAGGTATTAACACTTGAGATATGTTGGGGAGACTTTCACCTGCAATT 789
QY 121 CAGCAGTCCAAAGATGATTTTAAATACCAAAATTAAGTTGCAATTTCCGGGTAC 180
Db 790 CAGAGTCTAACCAGAGACCTTTGCTAGTCCAAATTCAGTCAAGAGTAAATGTTCC 849
QY 181 TTGTGACGNTGAGCAATGTTGGCAGCGATCTCCAGCTTGGCGATGTTGTGCA 240
Db 850 AAATTCAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 909
QY 241 TGCAGTGTGGCCATCTC 260
Db 910 TGGCAGCTCCACCATCTC 929

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/147,208
FILING DATE: 02-MAR-1999
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Rudolph, John R.
REGISTRATION NUMBER: 38,003
REFERENCE/DOCKET NUMBER: 7841-76
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1879 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

```

```

Query Match      15.1%: Score 53.8; DB 4; Length 1879;
Best Local Similarity 50.0%: Pred. No. 5.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0

OY      1 GCACGATTCATCCCATCTCTGAGGCGTCCGCCGCAATTAACAGTGGGAGTCCTCT 60
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      670 GCAAATTCATATATTTGAGGGAGAAATGCGCAGAAATTAGTCAACCCGAGATCT 729

OY      61 CCACCAAAATGATGACATGCTCGAGCTGGAGACAGATGGGGGTGCAATCCACCAAGTC 120
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      730 GCACCCAGATCCTAGCCCTAATTAACACTTGAGATTAAGTTGGGGAGACTTTCCACTGCAATT 789

OY      121 CACAGACCAAGGATGGCATTTTAAATACCCAAATAGATTCGAAATTTCCGCCGGAAC 180
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      790 CAAAGCTCTACCAAGSAGACCTTTGCTAGTCCAAATCAACGCAAAAGACGTAATGTTCC 849

OY      181 TTGTGACGNTGAGCAATGTTGCGCAGCGATCTCCAGCTTGCGCATATGTTGTTGAA 240
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      850 AAATTCAGTGTGACATCGATGATATATTAATCCCTATCATAGCTCATGTTGATAGA 909

OY      241 TGCAGTGTGGGCAATTCCT 250
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      910 TGGCACCCTCCACCATCTGC 929

RESULT 15
US-08-907-166-1/c
Sequence 1, Application US/08907166
Patent No. 5948666
GENERAL INFORMATION:
APPLICANT: Callen, Walter
APPLICANT: Mather, Eric
TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
FILE REFERENCE: 09010/027001
CURRENT APPLICATION NUMBER: US/08/907,166
CURRENT FILING DATE: 1997-08-06
NUMBER OF SEQ. ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2607
TYPE: DNA

```

```

; ORGANISM: Ammonifex degensii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2604)
US-08-907-166-1

```

```

Query Match      8.9%; Score 31.6; DB 2; Length 2607;
Best Local Similarity 60.5%; Pred. No. 0.43;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 202 CGCGACGTGATCTCCAGCTTGCGCATCATGTTGTCGATGCAAGTGTGCGCCATTCTCC 261
    |||| | | ||||| |||| | ||||| || ||||| |||||
DB 821 CGCGAGAGAGACCTCCAGCTGCTAAATGTTGGTCCCGGCCAGGGCGGATGATCTCTCC 762
    |||| | | ||||| |||| | ||||| || ||||| |||||
QY 262 TCTCTGACCAACCTTCGCGCGTGT 287
    ||| | | | |||| | | |
DB 761 TCTTCACGCGGCACATCGCGCGGAT 736
    ||| | | | |||| | | |

```

Search completed: March 26, 2003, 19:42:57
 Job time : 32.5276 secs